National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-55544-0

Canada

THE UNIVERSITY OF ALBERTA

THE VALIDITY OF CONVENIENT OBESITY INDICATORS

BY

J. DRU MARSHALL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY.

DEPARTMENT OF PHYSICAL EDUCATION AND SPORT STUDIES

EDMONTON, ALBERTA

FALL, 1989

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: J. Dru Marshall

TITLE OF THESIS: The Validity of Convenient Obesity Indicators

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE GRANTED: 1989

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

J. Dru Marshall

(Student's signature)

#303 10140 120th Street
Edmonton, Alberta
T5K 1Z8

DATE: OCTOBER 6, 1989

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled The Validity of Convenient Obesity Indicators submitted by J. Dru Marshall in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

.....
(Co-Supervisor)

.....
(Co-Supervisor)

.....

.....

.....

Date: OCTOBER 5, 1989

ABSTRACT

Convenient obesity indicators are widely used to establish prevalence statistics and as screening tools for obesity treatment programs; however, these indicators have been insufficiently validated. This research was conducted to examine the validity of convenient obesity indicators used in children aged 7 to 16 years.

Two studies were undertaken. The first study (N=533) examined the validity of 4 convenient indicators that had commonly accepted obesity classification points (relative weight, relative BMI, CSTF sum of skinfolds, and the O-scale) in reference to an "operational" standard for obesity that had been derived from factor scores composed of the common variance of all convenient measures studied used in conjunction with a visual rating. The diagnostic utility of the 4 indicators varied; the use of most indicators was gender-dependent and it was recommended that indicators should be chosen with reference to the purpose for classifying subjects as obese. Depending on the indicator used, obesity prevalence rates in this sample varied between 4.9 and 18.8%, sensitivities between 59.1 and 100.0% and specificities between 88.6 and 100.0%.

In the second study (N=540) a more commonly accepted gold standard for obesity in children was used (≥25%, ≥30% densitometrically-determined fat in males and females, respectively) as a reference to determine the diagnostic utility of 4 convenient indicators (relative weight, relative BMI, CSTF sum of skinfolds, and the triceps skinfold). Again, the diagnostic utility of the 4 indicators varied. Depending

on the indicator used, obesity prevalence rates for this sample varied between 5.4 and 20.7%, sensitivities between 64.0 and 100.0%, and specificities between 82.0 and 93.0%. There were inconsistencies in the utility of indicators that were common to both studies: this was probably a result of the use of different gold standards in each study.

The lack of a clearly accepted criterion measure for obesity is the key methodological shortcoming in the validation of convenient obesity indicators, and therefore, future research should be directed towards identifying a "gold standard" measure of obesity. Until identified, researchers should be aware of the inherent errors associated with the use of convenient measures and the potential for inappropriate labelling and stigmatizing.

ACKNOWLEDGEMENTS

I would like to thank all of the members of my committee who were invaluable to me during my tenure as a graduate student: my co-supervisor, Dr. Art Quinney, who gave me the freedom I needed to grow but who also knew when to step in with timely advice and support; my co-supervisor Dr. Murray Smith, who helped me believe in myself and who was always there to pull me out when the water got too deep; Dr. Donald Spady, whose constant guidance and assistance provided comfort and strength; Dr. Jane Watkinson, who always forced me to broaden my perspective and who served as an excellent role model; and Dr. Janice Johnston, who made me appreciate the finer points of research.

I would also like to thank my external examiner, Dr. Don Bailey, for his valuable time and expert advice; and Dr. Clark Hazlett, for his guidance, advice, and support. Thanks are also extended to the following special people: Merl, Gary, and Paul, without whom the first study would never have been completed; the staff and students of both McKernan and Kenilworth Junior High Schools; Pat Conger and the research assistants who collected the data in 1981; my fellow graduate students, particularly Gord, Pat, Brenda, and Garry; and especially to those friends who have always believed in me, even when I did not (Mom, Dad, Leah, Peter, Liz, and Pat).

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
	Bibliography	5
II.	A COMPARISON OF CONVENIENT INDICATORS OF OBESITY	7
	Bibliography	29
III.	THE VALIDITY OF CONVENIENT INDICATORS OF OBESITY	31
	Bibliography	57
IV.	GENERAL DISCUSSION AND CONCLUSIONS	60
	Bibliography	73
	APPENDIX 1:Review of Related Literature	75

LIST OF TABLES

TABLE	DESCRIPTION	PAGE
II.1	Summary of Available Measures	23
II.2	Frequencies and Percentages of the Sample Classified as Obese by the Obesity Indicators	24
II.3	Z Values for Tests of Differences Between Proportions Classified as Obese and Non-obese	25
II.4	Principal Components Solutions for Total Sample, Male, Female, and each School	26
II.5	Correlations between Adiposity Measures and Criterion of Visual Inspection	27
II.6	Diagnostic Capabilities of Four Adiposity Scales in Identifying Obesity	28
III.1	Age- and Gender-Specific Equations to Convert Body Density to Percent Body Fat	49
III.2	Age and Gender Distribution of the Sample	50
III.3	Summary of Available Measures	51
III.4	Frequencies and Percentages of the Sample Classified as Obese by the Obesity Indicators	52
III.5	Z Values for Tests of Differences Between Proportions Classified as Obese and Non-obese	53
III.6	Diagnostic Utilities of Four Adiposity Indicators in Identifying Obesity	54
III.7	Principal Components Solutions for Total Sample and Subsamples	55
III.8	Correlations Between Adiposity Indicators and Body Density Criterion for each Principal Components Solution	56
IV.1	Proportion of Samples Classified as Obese by Various Indicators	71

TABLE		PAGE
IV.2	Percent of Body Fat Cut-off Points at Various Percentiles in Males and Females (Study 2)	72
A.1	Body Density Values in Boys	135
A.2	FFB Density Values in Boys	136
A.3	Body Density Values in Girls	137
A.4	FFB Density Values in Girls	138
A.5	Acceptable Classification Points for Obesity Using Convenient Indicators	139

I. Introduction

Obesity, a common metabolic and nutritional disorder, has been defined as an excess of body fat leading to a significant impairment of health (1). This definition leads to two important questions: how is fat accurately measured?; and what amount of fat is considered excessive? Both questions are problematic, and thus the definition of obesity is statistically based and somewhat arbitrary. Many obesity indicators have more than one cut-off point (1,2), suggesting that the definition is value-laden rather than objective. The need for a quantifiable and generally accepted definition of obesity is paramount, and its availability basic to the accurate establishment of prevalence statistics and the organization of obesity treatment programs.

The measurement of body composition is an inexact science. All body composition measurement techniques provide indirect estimations for which accuracy is affected by measurement error and biological variability (3,4,5). Even cadaver analysis, which appears to be a direct measurement technique, has limitations in precise dissection of fat stores. Several indirect laboratory techniques have been developed, based on a variety of theoretical relationships and the use of "constants" derived from human and animal cadaver dissections (5). The three standard laboratory techniques, densitometry, total body water, and total body potassium, against which many laboratory techniques and all clinically convenient techniques are judged, use the two compartment model (5,6,7) to derive estimates of fat and lean tissue. The primary limitation of this model is the assumed constant density of the fat free body mass, a "constant" which varies with age, gender,

disease, and activity level (3,4,8,9). Thus, "constants" are not always generalizable to the population of interest. Most laboratory techniques are associated with a minimum error of 3 to 4% body fat (4,5), are expensive, clinically inconvenient or unavailable, and in some cases require specialized equipment, operator training, and some potential subject discomfort.

Because of these disadvantages, more indirect but clinically convenient and noninvasive measures of adiposity have been developed. These include various skinfold indices, taken singly or together (10,11,12), and various height-weight indices, such as relative weight, body mass index (BMI), and ponderal index. These indices have the advantage of speed and ease of use, but are "doubly indirect" methods and therefore can only be as accurate as the criterion methods against which they are judged.

For the definition of obesity, in addition to the problem of measuring body composition accurately, there is a problem with dichotomizing a continuous variable (ie., adiposity) into nonobese and obese categories. Fatness is a matter of degree. The ideal definition of obesity would involve the identification of the amount of fat that confers excess health risk. Unfortunately, the types of studies required to derive this definition have not been done. The point at which the body fat stores become clinically hazardous is unknown and is a subject of considerable debate (10,13,14,15). The relationship between body weight or BMI and mortality and morbidity has been described as linear, J-shaped, and U-shaped (16,17,18). In adults, marked obesity has adverse effects on health and longevity (17,19,20,21). The health risks of obesity in children are poorly

established (22), but obesity in childhood predisposes to obesity in adulthood, where the risks are evident (19,23). Therefore, since the exact level of increased fatness which confers excess health risk is unknown, definitions of obesity are statistically and somewhat arbitrarily derived (2,20,23).

A number of convenient adiposity indicators have cut-off points for obesity. These arbitrarily determined definitions may potentially be associated with errors that are costly to both the health care system and the individual. For example, false positive results may place an added burden on the health care system in terms of unnecessary and costly medical investigations. For the individual, a false positive result may end in unnecessary treatments that are expensive and potentially more damaging, both psychologically and metabolically, than the obesity itself (25,26). Conversely, a false negative result may aid in the development of serious medical conditions, such as diabetes or cancer (18,19,20), that are more costly to both the health care system and the individual than is obesity. Therefore, selection of a convenient obesity indicator should be made with a knowledge of the types of errors associated with each indicator.

Despite the aforementioned problems, convenient obesity indicators are widely used to establish prevalence statistics and as screening tools for obesity treatment programs, particularly in children and adolescents, where it is believed that early identification and treatment may help to alleviate future health problems (18). However, validation work with these indicators has been minimal (10). An adequate validating criterion method for adiposity, and subsequently obesity, is lacking. Recommended cut-off points for various indicators

have resulted in different proportions of the population being identified as obese (27,28,29), and thus confusing prevalence statistics have been reported. Few studies have been undertaken to compare the utility of different convenient indicators; those that have have utilized relatively small sample sizes (30,31,32), and therefore concerns related to statistical power, normative accuracy and generalizability remain. It was in reference to these problems that this research was conducted to evaluate the validity of convenient obesity indicators in children and adolescents. The primary objectives were to:

1. compare the utility of convenient adiposity indicators to classify individuals as obese or nonobese;
2. examine the diagnostic utilities of convenient obesity indicators in comparison to an "operational" standard of obesity and a "gold" standard of obesity; and
3. determine which indicator, if any, was suitable for use in children aged 7 to 14 years.

Bibliography

1. Burton BT, Foster WR, Hirsch J, Van Itallie TB. Health implications of obesity: an NIH Consensus Development Conference. Int J Obesity 1985;9:155-169.

2. Paige DM. Obesity in childhood and adolescence. Special problems in diagnosis and treatment. Postgrad Med 1986;79(1):233-245.

3. Brodie DA. Techniques of measurement of body composition. Part I. Sports Med 1988;5:11-40.

4. Brodie DA. Techniques of measurement of body composition. Part II. Sports Med 1988;5:74-98.

5. Lohman TG. Applicability of body composition techniques and constants for children and youths. In: Pandolf K, ed. Exercise and Sports Sciences Reviews 1986;14:325-357.

6. Keys A, Brozek J. Body fat in adult man. Physical Rev 1953;33:245-325.

7. Lukaski HC. Methods for the assessment of human body composition: traditional and new. Am J Clin Nutr 1987;46:537-556.

8. Boileau RA, Wilmore JH, Lohman TG, Slaughter MH, Riner WF. Estimation of body density from skinfold thicknesses, body circumferences and skeletal widths in boys aged 8 to 11 years: comparison of two samples. Hum Biol 1981;53(4):575-592.

9. Boileau RA, Lohman TG, Slaughter MH, Ball TE, Going SB, Hendrix MK. Hydration of the fat-free body in children during maturation. Hum Biol 1984;56(4):651-666.

10. Bandini LG, Dietz WH. Assessment of body fatness in childhood obesity: evaluation of laboratory and anthropometric techniques. Am Dietit Assoc 1987;87(10):1344-1348.

11. Canadian Standardized Test of Fitness (CSTF) Operations Manual. Third Edition. Ottawa: Minister of State, Fitness and Amateur Sport, 1986.

12. Pollock ML, Jackson AS. Research progress in validation of clinical methods of assessing body composition. Med Sci Sport Exer 1984;16(6):606-613.

13. Brennan N, Kevany J. Anthropometry and body image in a selected sample of adolescent girls. IJMS 1985;154(6):220-227.

14. Callaway W. Weight standards: their clinical significance. Ann Intern Med 1984;100(2):296-298.

15. Stark O, Lloyd JK. Some aspects of obesity in childhood. Postgrad Med Journal 1986;62:87-92.

16. Bray GA. Complications of obesity. Ann Intern Med 1985;103(6 pt 2):1052-1062.

17. Manson JE, Stampfer MJ, Hennekens CH, Willett WC. Body weight and longevity. A reassessment. JAMA 1987;257(3):353-358.

18. Simopoulos AP. The health implications of overweight and obesity. Nutr Reviews 1985;43(2):33-40.

19. Bray GA. Overweight is risking fate. Definition, classification, prevalence, and risks. Ann NY Aca Sci 1987;499:14-28.

20. Simopoulos AP. Characteristics of obesity: an overview. Ann NY Aca Sci 1987;499;4-13.

21. Simopoulos AP, Van Itallie TB. Body weight, health, and longevity. Ann Intern Med 1984;100:285-295.

22. Johnston FE. Health implications of childhood obesity. Ann Intern Med 1985;103(6 pt 2):1068-1072.

23. Dietz WH. Obesity in infants, children, and adolescents in the United States. I. Identification, natural history, and aftereffects. Nutr Research 1981;1:117-137.

24. Van Itallie TB. Health implications of overweight and obesity in the United States. Ann Intern Med 1985;103(6 pt 2):983-988.

25. Wooley SC, Wooley OW. Should obesity be treated at all? In: Stunkard AJ, Stellar E (eds.) Eating and Its Disorders. New York:Raven Press. 1984;185-192.

26. Brownell KD. The psychology and physiology of obesity: Implications for screening and treatment. J Am Dietit Assoc 1984;84(4):406-414.

27. Dietz WH. Childhood obesity: susceptibility, cause and management. J Pediatrics 1983;103:676-685.

28. Franklin B, Rubenfire M. Losing weight through exercise. JAMA 1980;244(4):377-379.

29. Hager A. Nutritional problems in adolescence - obesity. Nutr Rev 1981;39(2):89-95.

30. Roche A, Siervogel R, Chumlea W, Webb P. Grading body fatness from limited anthropometric data. Am J Clin Nutr 1981;34:2831-2838.

31. Rolland-Cachera MF, Sempe M, Guilloud-Bataille M, Patois E, Pequinot-Guggenbuhl F, Fautraud V. Adiposity indices in children. Am J Clin Nutr 1982;36:178-184.

32. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. J Adolesc Health Care. 1985;6(3):185-190.

Chapter II: A Comparison of Convenient Indicators of Obesity[1]

Obesity, a common metabolic and nutritional disorder, is frequently defined as an excess of body fat (1,2,3). However, the accurate measurement of body fat in a living organism is difficult. Several indirect laboratory measures of adiposity have been developed. Most of these methods are based on theoretical relationships and in many cases involve the use of "constants" that are not always generalizable to one's population of interest (4). These methods include: densitometry; the measurement of total body potassium, total body water, creatinine and 3-methylhistidine excretion; ultrasound, computed tomography, nuclear magnetic resonance imaging, infrared interactance, electrical impedance, and total body electrical conductivity (5,6). While these methods provide moderately accurate estimates of body fat and have been proposed in some cases as "gold standards", none has been proven to be both theoretically and empirically an unquestioned validating criterion method for the estimation of body fat. These laboratory methods are also expensive, clinically inconvenient or unavailable, time-consuming, and some require a high degree of operator training and subject cooperation. Consequently, clinically convenient but even more indirect, non-invasive measures of adiposity have been developed. Examples of these include: relative weight, body mass index (BMI) (wt/ht^2), relative BMI (1), ponderal index (PI) ($wt^{.33}/ht$), and various skinfold measures, taken singly or together.

---

[1] A version of this chapter has been accepted for publication. Marshall JD, Hazlett CB, Spady DW, Quinney HA. A Comparison of Convenient Indicators of Obesity. Am J Clin Nutr. In Press.

Several problems exist with these latter measures which prohibit each from being considered as sufficiently validated (7). The lack of an adequate validating criterion is the key methodological shortcoming in the development of any alternate laboratory or clinical measure of adiposity, and subsequently obesity. Recommended cut-off points for each clinical measure for purposes of classifying subjects as obese (or nonobese) have resulted in varying proportions of the population being so identified (2,8,9,10). Unfortunately few studies have been undertaken which incorporate all such measures in a common group, thereby enabling comparable obesity cut-off points to be derived for each measure. Comparative one sample studies using children and adolescents have been employed, but relatively small sample sizes were utilized (3,11,12), thus concerns related to statistical power, normative accuracy and generalizability remain.

The visual rating of adiposity, which represents a departure from traditional anthropometric field techniques, has shown promise. Trained observers have been visually able to rate percent body fat as accurately as that predicted from skinfold measurements (13). Visual ratings of adiposity are simple, quick, and inexpensive and may provide an alternative to some of the more costly techniques currently available.

At present, therefore, clinically convenient measures of adiposity, and subsequently obesity, have not been sufficiently validated. This study was undertaken to help advance research aimed at improving the level of accuracy in measuring obesity. It was assumed that since clinically convenient indicators of adiposity and obesity are widely used they possess some degree of validity as adiposity measures, even

though some inconsistencies are known to exist across measures. Given this, it was then also assumed that a measure derived from the common (or shared) variance of all convenient indicators was likely to be as or more valid than any single indicator. Since the classification point for obesity has not been uniformly agreed to, it is apparent that the level of increased fatness which causes increased health risk is unknown and therefore the present definition of obesity is more appropriately characterized as a statistical, value-laden judgement rather than an accurately measured pathological phenomonon. As such, obesity in this study was treated as a non-operationalized construct which, given its associated value-laden attributes, could also be assumed to be measured visually. Given these three key assumptions, the objectives of the research were:

1. to compare some convenient measures of adiposity, particularly with reference to their commonly derived classification points of obesity: relative weight, relative body mass index (BMI), the sum of 5 skinfolds (14) and the O-scale rating (15);
2. to determine (via principal components analysis) if a common underlying construct was evident among these various measures; and if so,
3. to determine if that construct was adiposity (by correlating the associated factor scores with a selected criterion visual inspection rating); and, if so,
4. to construct a standard for obesity classification from the relationship between the criterion visual inspection ratings and the derived factor scores; and
5. to determine, by comparison with the constructed standard, which,

if any, of these adiposity measures were better suited to correctly classifying obesity in children at the Junior High School level.

Subjects and Methods

Subjects (N=533) were recruited from two Junior High Schools in Edmonton, Alberta, Canada, a metropolitan center of some 600,000 people. Subjects were selected on the basis of convenience and obtained permission. Measures were taken over a two month period on children in grades 7, 8 and 9. Eligible students were those receiving parental consent and who personally agreed to be tested. A total of 293 out of a possible 321 (91.3%) students were tested in school 1 (156 males, 137 females; age range = 11.8 - 15.9 years), while 240 out of 258 (93.0%) students were tested in school 2 (124 males, 116 females; age range = 12.0 - 15.6 years). Subjectively, no apparent differences existed between tested and non-tested students. All aspects of this research were reviewed and approved by a medical ethics review panel.

Measurements

Anthropometric measurements taken included height, weight, and skinfold thicknesses. The criterion measure involved a four-point subjective visual rating of the subject's adiposity level. All measures were taken during physical education classes by a team of four trained observers. For purposes of efficiency, two of the four trained observers measured only height and weight, and the other two measured only skinfolds.

Prior to the study, intra- and inter-observer reliability was established on a volunteer sample of 30 subjects. Estimates established via the dependent analysis of variance model demonstrated that unadjusted reliabilities for all anthropometric measures ranged

from 0.96 (abdominal skinfold) to 1.00 (weight).

Height was measured to the nearest 0.2 cm with the shoes removed and the subject in indoor clothing. Weight was measured on a balance beam scale to the nearest 0.2 kg. The scale was calibrated with a set of known weights and any necessary adjustments made at the beginning of each testing day.

Triceps, subscapular, biceps, suprailiac, calf (14), supraspinale, abdominal and front thigh (15) skinfolds were measured to the nearest 0.2 mm using a Harpenden skinfold caliper following the general procedure outlined in the Canadian Standardized Test of Fitness Manual (14) and the O-scale manual (15).

Six adiposity measures were calculated from these measures of height, weight, and skinfolds, including three height-weight indices (relative weight, relative BMI, and PI), and three skinfold indices (triceps skinfold, Canadian Standardized Test of Fitness sum of skinfolds (14), and the O-scale geometrically scaled sum of skinfolds (15)) For the purposes of this study PI was calculated as $ht/wt^{.33}$, as is commonly done in the Sheldon somatotyping system (16).

The National Center for Health Statistics percentiles (17) were used for the calculations of relative weight and relative BMI.

Relative weight (18) was obtained by:

$$\frac{\text{subject weight}}{\text{expected weight*}} \times 100$$

* for height, age, and gender at the 50th percentile

Relative BMI (1) was obtained by:

$$\frac{\text{actual BMI}}{\text{BMI 50 *}} \times 100$$

*BMI at the 50th percentiles of height and weight for same age and gender.

The CSTF skinfold test (14) involved the addition of five skinfolds: triceps, biceps, subscapular, suprailiac, and medial calf.

The O-scale system involved the addition of six skinfolds (triceps, subscapular, supraspinale, abdominal, front thigh, and medial calf). This sum was then geometrically scaled to a standard stature of 170.18 cm (sum of skinfolds x 170.18/ht(cm)) and the scaled value was compared to normative data to derive a final adiposity rating ranging from 1 to 9 (15).

A 4-point visual inspection rating scale was developed and was used by each rater to subjectively classify adiposity prior to the collection of all other measures. A rating of 1 corresponded to "slim" (thin, anorexic-like), 2 to "ideal" (optimal weight to height), 3 to "overweight" (pleasantly plump, but not indicative of health risk), and 4 to "obese" (grossly overweight). An inter-rater reliability study of this developed scale, utilizing 75 subjects (40 males, 35 females), established an unadjusted reliability estimate of 0.95, indicating that an unusually high degree of consistency across raters was attainable by simple visual inspection.

Obesity Classifications

Four of the six adiposity measures (relative weight, relative BMI, CSTF sum, O-scale rating) had commonly accepted obesity classification points. A value greater than or equal to 120% represented obesity for both the relative weight (19) and relative BMI (1) measures. For the CSTF skinfolds, a sum greater than or equal to the value corresponding to the normative eighty-fifth percentile for age and gender was defined as obese (20), while an adiposity rating of 8 or 9 for the O-scale system was so chosen.

Analyses

To determine if inconsistencies in classifying obese subjects occurred across measures, differences in correlated proportions tests (21) were undertaken. To determine if a common construct underlay the adiposity measures, a principal components analysis (22) was performed on six adiposity measures: three height-weight indices (relative weight, relative BMI, and PI), and three skinfold indices (triceps skinfold, CSTF sum, and O-scale sum). The balance of these two types of indices in the principal components analysis was intentional, facilitating factors that were either method or construct related to emerge in the solution. Derived factor scores were examined for linear and non-linear relationships using appropriate statistical models and data plots. Based on these statistical and visual inspections, an obesity classification point for the factor scores was determined. Utilizing dichotomized factor scores as the standard for obesity classification, the diagnostic utilities of the four adiposity measures that had generally accepted obesity cut-off points were examined. The diagnostic utilities examined included sensitivity (proportion of "truly" obese subjects detected by the measure), specificity (proportion of "truly" nonobese subjects detected), overall accuracy (proportion of correct classifications in sample), positive predictive value (the probability that a subject who has been measured as obese is "truly" obese), and negative predictive value (the probability that a subject who has been measured as nonobese is "truly" nonobese) (23).

Results

Although 533 subjects were recruited, some measurements were unavailable for certain subjects (e.g., a subject wearing a cast made

certain skinfold measures inaccessible and weight inaccurate). A summary of the measures taken, and the corresponding response rates can be found in Table II.1. All further analyses reported here were based on sample sizes where no missing values were present in any data involved in the respective analyses.

Consistency in Obesity Classifications

The proportion of the sample classified as obese using accepted standards varied from 8.4% (O-scale rating) to 18.8% (CSTF sum) (Table II.2). Statistical tests for differences among correlated proportions resulted in significant differences in obese classifications among measures (Table II.3), except for one comparison (relative weight and O-scale rating). Thus, in this sample the previously noted inconsistencies (2,8,9,10) across obesity screening measures was confirmed. The visual rating scale developed for this study also identified statistically different proportions of obesity, classifying the least number of subjects as obese (Table II.3).

Construct Validation

A principal components analysis of six measures of adiposity (relative weight, relative BMI, PI, triceps skinfold, CSTF sum of skinfolds, O scale rating) resulted in a one-factor solution[2] accounting for 85.6% of the variance (Table II.4). Factor loadings for each original adiposity measure ranged from 0.86 to 0.96. Principal components analyses for these same 6 indicators were repeated on subsamples of each sex and school and the solutions also resulted in a one-factor solution[2] (Table II.4). Since all solutions consistently

---

[2]Other than for the first factor, all eigenvalues were less than 1.0.

indicated these measures were unifactorial, evidence existed that only one main construct underlay all of the original six measures of adiposity.

The visual inspection rating for each subject was correlated with each of the original six adiposity measures and with the factor scores derived from the above total sample principal components solution (Table II.5). The correlations with visual inspection were all statistically significant (ranging from 0.66 to 0.81), with the factor scores having the strongest linear relationship (0.81).

The data were plotted to further investigate the nature of the relationship between the visual inspection ratings and the derived factor scores. The relationship was obviously linear and the selection of an obesity classification point of greater than or equal to 1.5 standard deviations above the mean of the factor scores resulted in noteworthy consistencies. One hundred percent of the subjects who had been visually rated as "obese", and 100% of all those visually classified as "slim" or "ideal" were correctly classified if greater than or equal to 1.5 standard deviations above the mean of the factor scores was defined as the obesity classification point. The only inconsistencies evident were among those visually labelled as overweight: i.e., 3.8% of individuals visually classified as overweight were classified as obese according to the dichotomized factor score.

At face value therefore, the factor solution appeared to be not only unifactorial, but 1) the underlying construct was logically that of adiposity, and 2) an imposed classification point for purposes of diagnosing or labelling some subjects as extreme (i.e., obese), appeared to be more meaningfully set at approximately the 92nd

percentile point[3] on a measure of adiposity (in contrast to percentile points previously used).

Given these findings, a factor score of greater than or equal to 1.5 was proposed as a standard for the classification of obesity in this sample of junior high school students. This proposition enabled further examination of the original adiposity measures for which accepted obesity classification points were available, through an analysis of the sensitivity, specificity, overall accuracy, and predictive value of each measure (Table II.6).

Comparisons among these four measures demonstrated that the utility of these scales varied, particularly in reference to the sex of the subject. Examination of Table II.6 revealed that all scales for both genders met acceptable standards for negative predictive values (NPV)- i.e., if a scale identified a child as nonobese, the probability is high (0.96 to 1.0) that the child is nonobese. However, since prevalence of obesity is assumed to be only 8%, all scales should have acceptable NPVs. Given this, the utility of each scale could be better distinguished by an examination of the sensitivity, specificity and positive predictive value (PPV) of each scale.

Regardless of sex, relative BMI was found to have 100% sensitivity; however, it is apparent that in identifying true obesity so accurately, some nonobese subjects were "incorrectly" labelled as obese, as indicated by the relatively low specificity (93.0). This is confirmed by an examination of the PPV. That is, if the relative BMI measure

---

[3] The adiposity distribution in this sample was tested for normality and rejected at the .05 level. The distribution was positively skewed, with $\geq$ 1.5 SD corresponding to the 92 percentile point.

indicated a child was obese, there is only a 0.55 probability that the child is truly obese (PPV).

The CSTF was less sensitive overall (92.9%) than the relative BMI measure, but this reduced sensitivity was restricted to females only (86.4%), males being 100%. Although the CSTF was very sensitive for males, it was not very specific (88.6%), resulting in an unacceptable accuracy in the diagnostic ability of CSTF for obesity (PPV = 0.43). For females, however, the CSTF was highly specific (98.7%), resulting in an acceptable PPV of 0.88.

The utility of the O-scale also varied with gender. This scale identified all true obese male subjects (100% sensitivity), but in doing so also classified some nonobese as obese (96.5% specificity), resulting in a PPV value of 0.71. When the O-scale is used with female subjects, only 59.1% of the all obese females will be so identified. However, given the apparent high cut-off point for obesity in females of this scale, any female identified as obese has a 1.0 probability of being truly obese (PPV).

The relative weight measure was found to have reasonable utility in measuring obesity in male subjects (90% sensitivity; 96.9% specificity; 0.72 PPV). Its apparent weakness for measuring obesity in females is evident when one considers that two of every three obese females would be correctly labelled as obese (sensitivity of 68.2%), and that only 55% of the time a female subject would have a true obesity diagnosis (PPV = 0.55).

Discussion

An examination of Table II.3 confirmed that the various convenient adiposity measures which have been developed and used for screening

obese subjects are disturbingly inconsistent in identifying subjects with this disorder. In this regard, this study's results are similar to those findings of other authors who have reported variations of 3 to 25% of children and teenagers classified as obese, depending on the measure used (2,8,9,10). It is important to note that inconsistencies were evident within height-weight indices, within various skinfold indices, between both groups of measures, and among all of these with a subjective visual rating. This latter finding also supported that of Moynihan et al (24), who noted inconsistencies between visual ratings and relative weight. However, the presence of such inconsistencies is logical since previously recommended obesity classification points in these various scales have not been based on a common percentile point. Obviously such would be needed to ensure reliability in classification, and as a prerequisite to a valid classification of obesity.

The principal components analyses consistently indicated a unidimensional solution, despite the fact an equivalent number of height-weight indices and skinfold indices were analyzed. This is in contrast to Mueller and Wohlleb (25) who found a two factor solution (based on only skinfold measures), interpreted to be an obesity index and a trunk-extremity distribution of fat index. However, their conclusions were predicated on use of factors that had associated eigenvalues that were less than one. The retention and importance of factors which account for even less variance than one variable that has been entered into a factor solution is statistically illogical and of dubious value clinically. In this study, selected measures were more heterogeneous, both methodologically and anthropometrically, than those reported by Mueller and Wohlleb. Since all these measures have been

regarded as legitimate convenient measures of adiposity, it appeared reasonable to assume that the primary influence on variance across subjects was likely adiposity, rather than any method- or anatomically-based attribute.

The criterion measure of visual inspection was highly correlated with the factor scores, providing logical support that the factor scores were measures of adiposity. However, no previous studies have used visual inspection as a criterion measure. In fact, visual inspection has been shown to classify obesity differently than relative weight (24), and the results of this study similarly demonstrated such inconsistencies (Table II.3). However, such inconsistencies in the absence of any gold standard do not necessarily imply visual inspection to be an inadequate criterion. Sterner and Burke (13) demonstrated that some observers can visually estimate percent fat almost as accurately as it can be calculated from skinfold measurements when densitometry is used as the criterion method.

Further, visual inspection does possess inherent face validity for measuring obesity, if not adiposity. Since the nonobese and obese dichotomies of visual inspection and factor scores resulted in an overall consistency of 96.2%, it was hypothesized that the underlying construct of the factor solution was adiposity. Therefore, factor scores reflecting only the common variance across all measures of adiposity were seen to be more ideal than any one of the individual measures per se. Acceptance of this logic is prerequisite to the examination of the apparent ramifications one encounters in utilizing relative weight, CSTF, O-Scale, or relative BMI for classifying obesity in junior high school children. The results of this study support use

of some of these scales, but only after some consideration of gender, and in reference to one's purpose for identifying obesity in children.

It is important to identify obese children for treatment and prevention of health risk. However, treatment is unfortunately rarely successful, and the inclusion in treatment and labelling of "obesity" in some children results in negative psychological and social consequences that may be potentially worse than the effects of obesity itself (26).

One scale (relative BMI) did not vary significantly by gender. Viewed in terms of a priori probability, it possessed high sensitivity (100%), specificity (93%), and overall accuracy (93.5%). At the clinical level, however, the low PPV of this test (0.55) indicates that although all truly obese individuals are captured, there is only a 0.55 probability that a child so identified will be actually obese. Thus, if one is interested in finding all true obese children, the relative BMI is the measure of choice, but its use has associated potential psychological-social costs since some nonobese children are identified as obese.

The utility of other scales varies by gender. For males, the common obesity classifications derived from the relative weight and O-scale measures provide relatively high sensitivities, specificities, accuracies, and PPVs. Thus, there is relatively little psychological-social or clinical cost incurred in using these measures with males. There is a psychological-social cost, however, when the CSTF is used with males. Although the CSTF identifies true obesity, it does so at the cost of lowered specificity, and therefore some nonobese boys will be labelled as obese.

For females, the CSTF measure is better than the relative BMI measure, having reduced sensitivity, but higher specificity, accuracy, and PPV. Although the CSTF does not identify true obesity as well as the relative BMI measure, its use is associated with relatively less psychological-social costs. The O-scale and relative weight measures, however, both possess relatively low levels of sensitivity, and therefore the ability to detect true obesity in females with these scales is of concern. With the O-scale, some obese females will be classified as nonobese, but the PPV of this test demonstrates that if this scale identifies a subject as obese, she will be obese. The relative weight measure is less useful than the O-scale since, in addition to being insensitive, it is also lacking in clinical utility, having a PPV of 0.55.

Differences in the utility of convenient adiposity indices between genders have been previously noted (3,11,12). Johnston (12) found that relative body weight and triceps skinfold measures differed in accuracy between genders. Both were low in sensitivity (23-50%) and high in specificity (85-100%) when percent fat, as determined by underwater weighing, was used as the obesity criterion (males = 25%, females = 30%).

The difference in the utility of these measures by gender could be influenced by the earlier maturation of girls in comparison to boys. To test this hypothesis, an examination of older males who were in a period of rapid growth would be required. Unfortunately pubertal status was not assessed in this study and a larger sample would be necessary before an adequate subsample could be extracted for such specific analyses. Regardless, the findings suggest more research is required to

identify appropriate obesity classification points for each gender.

Pursuit of this methodology, using additional measures that are generally accepted as criterion measures of body fat, such as densitometry, is recommended. With such, the appropriate development of prevalence statistics for obesity can be more reliably identified. Until such time, it is suggested that the measurement of obesity in junior high students be done with some regard for 1) the degree of inherent error that various measures have, 2) the potential for inappropriate labelling and stigmatizing, and 3) the possibility that the proportion of children with this disorder may be less than the rates that some other studies have indicated.

Table II.1: Summary of Available Measures

Variable	Male $\bar{x} \pm$ SD (n)	Female $\bar{x} \pm$ SD (n)	Total Sample $\bar{x} \pm$ SD (n)
Triceps (mm)	11.2 ± 5.0 (280)	15.2 ± 4.8 (252)	13.1 ± 5.3 (532)
CSTF sum (mm)	49.5 ± 26.1 (279)	60.9 ± 21.3 (252)	54.9 ± 24.6 (531)
O-Scale sum (mm)	65.9 ± 33.1 (275)	83.4 ± 28.5 (252)	74.3 ± 32.2 (527)
Relative weight (%)	99.6 ± 17.6 (279)	99.8 ± 16.3 (253)	99.7 ± 17.0 (532)
Relative BMI (%)	102.0 ± 18.5 (279)	104.0 ± 16.6 (253)	103.0 ± 17.6 (532)
PI	13.3 ± 0.7 (279)	13.2 ± 0.6 (253)	13.2 ± 0.6 (532)
Visual Rating	2.0 ± 0.7 (280)	2.0 ± 0.7 (252)	2.0 ± 0.7 (532)

Table II.2: Frequencies and Percentages of the Sample Classified as Obese by the Obesity Indicators

INDICATOR	FEMALES		MALES		TOTAL SAMPLE	
	(N)	(%)	(N)	(%)	(N)	(%)
Relative Weight	25	9.9	30	10.8	55	10.4
Relative BMI	41	16.2	38	13.7	79	14.8
CSTF Sum	40	15.9	60	21.5	100	18.8
O-scale	13	5.2	32	11.5	45	8.5
Visual Inspection	10	4.0	16	5.7	26	4.9

Table II.3: Z Values for Tests of Differences Between Proportions Classified as Obese and Non-Obese

	Relative Weight	Relative BMI	CSTF Total	O-Scale Rating	Visual Inspection
Relative Weight					
Relative BMI	3.90*				
CSTF Total	5.42*	2.99*			
O-Scale Rating	-1.73	-5.25*	7.35*		
Visual Inspection	-4.90*	-7.28*	8.60*	-3.80*	

*Statistically significant ($alpha_2$ = .05)

Table II.4: Principal Components Solutions for Total Sample, Male, Female and each School

Indicators	FACTOR SOLUTIONS				
	Solution 1 Total Sample	Solution 2 Males Only	Solution 3 Females Only	Solution 4 School 1 Only	Solution 5 School 2 Only
Triceps	.89	.91	.90	.90	.89
CSTF sum	.96	.96	.96	.96	.95
O-scale total	.96	.96	.96	.96	.95
Relative weight	.86	.90	.83	.87	.84
Relative BMI	.94	.94	.94	.94	.93
PI	.94	.93	.95	.95	.94
% of variance	85.6%	87.3%	85.8%	86.6%	84.3%

Table II.5: Correlations between Adiposity Measures and Criterion of Visual Inspection

Measures	Visual Inspection r (n)
Triceps	.72 (532)*
CSTF sum	.79 (531)*
O-scale total	.78 (527)*
Relative weight	.66 (531)*
Relative BMI	.78 (531)*
PI	.78 (531)*
Factor scores	.81 (526)*

*statistically significant at $\alpha_2 = .05$

Table II.6: Diagnostic Capabilities of four adiposity scales in Identifying Obesity*

Diagnostic Capability		Relative Weight	CSTF Total	O-Scale Rating	Relative BMI
A.	Sensitivity				
	Overall	78.6	92.9	78.6	100.0
	Female	68.2	86.4	59.1	100.0
	Male	90.0	100.0	100.0	100.0
B.	Specificity				
	Overall	96.1	93.4	98.1	93.0
	Female	95.2	98.7	100.0	91.7
	Male	96.9	88.6	96.5	94.1
C.	Accuracy				
	Overall	94.7	93.3	96.6	93.5
	Female	92.9	97.6	96.4	92.5
	Male	96.4	89.4	96.7	94.5
D.	Positive Predictive Value				
	Overall	0.64	0.55	0.78	0.55
	Female	0.55	0.88	1.00	0.51
	Male	0.72	0.43	0.71	0.60
E.	Negative Predictive Value				
	Overall	0.98	0.99	0.98	1.00
	Female	0.97	0.99	0.96	1.00
	Male	0.99	1.00	1.00	1.00

*Based on accepted obesity classification points for each scale compared with a standard of a dichotomized factor score at $\geq$ 1.5 SD.

Bibliography

1. Poskitt E. Management of Obesity. Arch Dis Child 1987;62:305-310.

2. Hager A. Nutritional problems in adolescence-obesity. Nutr Rev 1981;39(2):89-95.

3. Roche A, Siervogel R, Chumlea W, Webb P. Grading body fatness from limited anthropometric data. Am J Clin Nutr 1981;34:2831-2838.

4. Lohman T. Applicability of body composition techniques and constants for children and youths. In: Pandolf K, eds. Exercise and Sports Sciences Reviews 1986;14:325-357.

5. Brodie DA. Techniques of measurement of body composition Part I. Sports Med 1988;5:11-40.

6. Brodie DA. Techniques of measurement of body composition Part II. Sports Med 1988;5:74-98.

7. Bandini L, Dietz W. Assessment of body fatness in childhood obesity: evaluation of laboratory and anthropometric techniques. J Am Dietet Assoc 1987; 87(10):1344-1348.

8. Dietz W. Childhood obesity: Susceptibility, cause, and management. J. Pediatrics 1983;103:676-685.

9. Franklin B, Rubenfire M. Losing weight through exercise. JAMA 1980;244(4):377-379.

10. Heston M. Childhood Obesity: Implications for physical education. Runner 1985;23(3):3-6.

11. Rolland-Cachera MF, Sempe M, Guilloud-Bataille M, Patois E, Pequinot-Guggenbuhl F, Fautraud V. Adiposity indices in children. Am J Clin Nutr 1982;36:178-184.

12. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. J Adolesc Health Care 1985;6(3):185-190.

13. Sterner TG, Burke EJ. Body fat assessment: A comparison of visual estimation and skinfold techniques. Phys & Sportsmed 1986;14(4):101-107.

14. Canadian Standardized Test of Fitness (CSTF) Operations Manual. Ottawa: Minister of State, Fitness and Amateur Sport, 1986.

15. Ross W, Ward R. The O-scale system. Vancouver:Rosscraft Ltd., 1984.

16. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. J Chron Dis 1972;25:329-343.

17. Hamill P, Drizd T, Johnson C, Reed R, Roche A, Moore W. Physical growth: National center for health statistics percentiles. Am J Clin Nutr 1979;32:607-629.

18. Schwartz M. Childhood Obesity. Surgical Clin North Am 1979;59(6):995-1005.

19. Richmond D, Blyler E, Linscheid T. The obese child. Am Fam Phys 1983;28:129-134.

20. Canada Fitness Survey: Canadian Youth Fitness. Ottawa: Fitness Canada, Fitness and Amateur Sport, 1985.

21. Gravetter FJ, Wallnau LB. Statistics for the Behavioural Sciences. New York: West Publishing Company, 1988.

22. Harman HH. Modern Factor Analysis. Chicago: The University of Chicago Press, 1967.

23. Fletcher RH, Fletcher SW, Wagner EH: Clinical Epidemiology - the essentials. London: Williams & Wilkins, 1982.

24. Moynihan C, Stark O, Peckham C. Obesity in 16-year olds assessed by relative weight and Doctor's rating. Int J Obesity 1986;10:3-10.

25. Mueller WH, Wohlleb JC. Anatomical distribution of subcutaneous fat and its description by multivariate methods: how valid are principal components? Am J Phys Anthrop 1981;54:25-35.

26. Brownell KD. The psychology and physiology of obesity: Implications for screening and treatment. J Am Dietit Assoc 1984;84(4):406-414.

Chapter III: The Validity of Convenient Indicators of Obesity[1]

Obesity is defined as an excess of body fat frequently resulting in a significant impairment of health (1). There are two problems with this definition: how is body fat accurately measured; and how much fat is too much? The need for a quantifiable definition of obesity is great, and its availability basic to the accurate establishment of prevalence statistics and the organization of treatment programs.

The valid measurement of body fat in the human is difficult. All methods of body composition measurement are considered to yield indirect estimations for which accuracy is affected by measurement error and biological variability. Standard, well-accepted laboratory procedures, such as densitometry and hydrometry, are expensive, require specialized equipment and operator training, and involve potential subject discomfort. They are associated with a minimum error of approximately 3 to 4% body fat (2,3). Clinically convenient indicators, such as skinfolds and height-weight indices, are validated against these laboratory "gold" standards and thus must be considered to be even more indirect and potentially suspect.

Obesity, like baldness, is a matter of degree and there is a problem with dichotomizing the continuous trait of adiposity into obese and nonobese categories. It is known that a certain amount of fat is critical to survival and that there is a curvilinear relationship between fatness and health risk (4,5); however, the point at which the

---

[1] A version of this chapter has been submitted for publication. Marshall JD, Hazlett CB, Spady DW, Quinney HA, Conger PR, Johnston JJ. Am J Clin Nutr.

body fat stores become clinically hazardous is unknown and is a subject of considerable debate (6,7,8). Therefore, definitions of obesity are statistical, value-laden judgements rather than statements clearly defining a pathological condition.

Despite these problems, clinically convenient indicators are used to establish obesity prevalence statistics and as screening tools for obesity treatment programs. However, the use of these indicators is as problematic as the definition of obesity. The measures have not been sufficiently validated: in most cases validation has been done with small sample sizes and therefore concerns about statistical power, normative accuracy, and generalizability remain (9,10,11). Common percentile points for obesity classification have not been developed and consequently indicators classify the same individual differently, resulting in confusing prevalence rates (6,12,13,14). In addition, similar classification criteria are often used for children, adults and for both genders, although there is little justification for this (15).

The accurate diagnosis of obesity in children and adolescents is particularly important. Inaccurate identification can cause psychological and sociological problems (6,7,16) and unnecessary medical investigation. Although the health risks of obesity in children are poorly established, obesity in childhood predisposes to obesity in adulthood (17), where the health risks of obesity are clear (18,19). Therefore, early accurate diagnosis and intervention is important if reduction of psychological trauma and later health problems is to occur.

In previous work (20), we examined the accuracy of convenient obesity indicators in Junior High School students, aged 12 to 16 years,

and demonstrated that although the studied indicators classified varying proportions of the sample as obese, all were composed of a common underlying construct, adiposity. A criterion measure for obesity was defined using the relationship between derived factor scores from a principal components analysis and scores from a visual rating scale. Using this "operational standard" for obesity, we examined the diagnostic utilities of 4 indicators: relative weight, relative body mass index (21), sum of 5 skinfolds (22), and the O-scale (23), for which accepted obesity classification points were available, and found that the utility of the four measures varied. We concluded that: 1. the use of most of these measures should be considered gender-dependent with respect to the early adolescent subject; and 2. the measures should not be used without reference to the purpose for classifying subjects as obese. Although the study provided some practical information, there was one primary limitation: the lack of an "accepted" gold standard criterion measure of obesity. Visual inspection in conjunction with factor scores had not been used as a criterion method of obesity before. Consequently, we have used similar methodology in the present study to examine the validity of convenient indicators of obesity, with the exception that we used an accepted standard for obesity: percent body fat as estimated from body density derived from hydrostatic weighing. The specific objectives were:

1. to compare the utility of the following convenient measures of adiposity in categorizing a person as being obese or non-obese: relative weight, relative body mass index (BMI), the sum of 5 skinfolds, and the triceps skinfold;
2. to determine which, if any, of these measures was better suited

to accurately classifying obesity in children aged 7 to 14 years, through a comparison of their diagnostic utilities, and;

3. to determine any inconsistencies in the diagnostic utilities of the convenient indicators that were present between our previous work, which utilized a criterion measure of visual inspection in conjunction with factor scores, and this study, which utilized a more widely recognized criterion, percent body fat.

Subjects and Methods

These data on 600 children and youth, aged 7-14 y, were originally collected in 1981 in Edmonton, Alberta to aid in the development of normative data for use in the Canada Fitness Survey (24). A two-stage stratified, probability sampling frame was implemented. In the first stage, the city was stratified into quadrants and schools were randomly selected proportional to the number of schools in each quadrant. Exclusions to this population included all private schools and schools for the handicapped. In the second stage, subjects were stratified by age, gender, and school; disproportionate sampling was undertaken such that 100 males and 100 females were included in each of the age ranges 7-9, 10-12, and 13-14 y. The number of subjects drawn was proportional to the school's relative size among those schools that had been randomly selected.

Informed consent was obtained for each subject and all were briefed and familiarized with the laboratory procedures prior to any measurements being taken. All measures were conducted in the physical education laboratory at the University of Alberta. All aspects of this research were reviewed and approved by an ethics review panel. Partial analyses of these subjects have been reported elsewhere (25).

Measurements

Height was measured to the nearest mm on a portable stadiometer and weight to the nearest 0.1 kg on a calibrated beam balance scale. Triceps, biceps, subscapular, suprailiac, and calf skinfolds were measured to the nearest 0.2 mm using a Harpenden skinfold caliper following the general procedure outlined in the Canadian Standardized Test of Fitness (CSTF) Manual (22).

Body density, estimated via hydrostatic weighing (26), served as the criterion measure for adiposity. Total lung volume was measured with the subject in the water employing a closed circuit helium dilution technique (27,28). A Godart Pulmotest (De-Bilt-Utr, Holland) and helium analyzer (Godart) were used for lung volume measurements and a calibrated load cell was used to measure submerged weight. Underwater weight was measured several times until results were consistent. Water temperature was maintained at $33 \pm 1^{o}C$.

Twenty percent of the original sample was randomly selected for a retest which occurred between 21 and 28 days following the original test day. Intra-examiner reliability coefficients ranged from 0.96 (suprailiac skinfold) to 0.99 (height). Reliability studies were not conducted on the hydrostatic weighing technique, but the standard errors obtained were congruent with data in the literature.

Five adiposity indicators were calculated from the measures of height, weight, and skinfolds, including 3 height-weight indices (relative weight, relative BMI, and ponderal index), and 2 skinfold indices (triceps skinfold, and the CSTF sum of skinfolds).

The National Center for Health Statistics percentiles (29) were used for the calculations of relative weight and relative BMI. Relative

weight was obtained by:

$$\frac{\text{subject weight}}{\text{expected weight*}} \times 100$$

*for height, age, and gender at the 50th percentile

BMI was calculated as weight in kilograms divided by height in meters squared. Relative BMI (21) was obtained by:

$$\frac{\text{actual BMI}}{\text{BMI 50*}} \times 100$$

*BMI at the 50th percentile of height and weight for same age and gender.

Ponderal index was calculated as height in inches divided by the cube root of weight in pounds (30).

The CSTF skinfold test (22) involved the addition of all five skinfolds: triceps, biceps, subscapular, suprailiac, and medial calf.

Percent body fat, predicted from body density, served as the criterion measure for obesity. Body density was converted to percent fat using Lohman's (2) age- and gender-specific equations, which account for the lower bone mineral and higher body water content of prepubertal and pubertal children (Table III.1).

Obesity Classifications

For the criterion definition of obesity in children, an individual was categorized as obese if the percent body fat predicted from body density was $\geq$ 25% in males or $\geq$ 30% in females (3,11,18). Four of the five adiposity indicators (relative weight, relative BMI, CSTF sum, triceps skinfold) had commonly accepted obesity classification points. A value of $\geq$ 120% represented obesity for both the relative weight (18) and relative BMI (21) measures. For the CSTF skinfolds, a sum greater than or equal to the value corresponding to the normative 85th percentile for age and gender for the Canada Fitness Survey was defined

as obese (24). A triceps skinfold value ≥ 85th percentile for age and gender was defined as obese (6). The sample was split into age groups of 7-8, 9-10, 11-12, and 13-14 years, and the 85th percentile cut-off points for the triceps skinfold were determined for each age group and gender.

Analyses

Dependent proportions tests (31) were done to determine if statistically significant inconsistencies existed across indicators in the classification of obese subjects. Utilizing the dichotomized percent body fat measures (in reference to obesity) as a gold standard, the diagnostic utilities of the four adiposity indicators that had generally accepted obesity cut-off points were examined. The diagnostic utilities examined included sensitivity (proportion of "truly" obese subjects detected by the measure), specificity (proportion of "truly" nonobese subjects detected), overall accuracy (proportion of correct classifications in sample), positive predictive value (the probability that a subject who has been measured as obese is "truly" obese), and negative predictive value (the probability that a subject who has been measured as nonobese is "truly" nonobese) (32). In cases where inconsistencies were found in the diagnostic utility of some indicators between our previous work and this study, further correlational analyses[2] were performed.

Principal components analyses (34) were performed on the five adiposity indicators to determine if a common construct underlay all

---

[2] Triceps skinfold and CSTF sum of skinfolds were transformed logarithmically ($\log_e$) as it has been established that the relationship between skinfolds and body density is curvilinear (2,33); this relationship was confirmed in this data set.

indicators. To determine the construct validity of the factor solutions, body density was correlated with the derived factor scores from the principal components analyses and with each individual indicator and these were tested for statistical significance from the null and among each other.

A subsample consisting of all subjects older than 11 years of age was similarly analyzed to allow for closer comparison with the previous study (20). This subsample consisted of 248 subjects: 124 males and 124 females. In addition, to investigate the potential of puberty as a confounding variable, genders were analyzed separately with the potential periods of rapid growth eliminated from the sample (females - aged 10-12 y, males - aged 12-14 y). All statistical tests were conducted at the α_2=.05 level. Subjects with incomplete data or who had not tolerated the underwater weighing procedure were excluded from the analyses.

Results

Although 600 subjects were recruited, complete data were available for only 540 individuals (90%). A description of the age and gender distribution of the sample can be found in Table III.2, and a summary of all measures taken in Table III.3.

Consistency in Obesity Classification

Previously noted inconsistencies (20) across obesity indicators were confirmed in this sample. The selected criterion measure, percent body fat, classified the least number of subjects as obese (5.4%), whereas the proportion of the sample classified as obese by the convenient indicators varied from 11.7% (relative weight) to 20.7% (CSTF sum) (Table III.4). Statistical tests for differences among

correlated proportions were all statistically significant (Table III.5) with the exception of only two comparisons (the triceps skinfold with relative BMI and relative weight).

Diagnostic Utilities

Comparison among the four indicators for which commonly accepted obesity classification points were available showed that the utility of these measures varied, particularly in reference to the gender of the subject (Table III.6). All scales for both genders met acceptable standards for negative predictive value (NPV), i.e., if an indicator identified a child as nonobese, the probability is high (0.98 - 1.00) that the child is nonobese. However, since the prevalence of obesity was identified to be only 5.4% by the gold standard of percent body fat, all indicators should have very high NPVs. As such, the utility of each indicator was better distinguished through an examination of the sensitivity, specificity, overall accuracy, and positive predictive value (PPV) of each indicator.

The utility of the CSTF sum and relative BMI indicators did not vary by gender. The CSTF sum of skinfolds was found to have 100% sensitivity; however, in identifying "true" obesity so accurately some nonobese subjects were classified as obese, as indicated by the reduced specificity (84%) and accuracy (85%) of the indicator. This is also confirmed by an examination of the PPV, which suggests that if the CSTF sum indicated a child was obese, there is only a 0.26 probability that the child is truly obese. The relative BMI was less sensitive (93%) than the CSTF sum, but slightly more specific (87%), indicating that although it did not identify "true" obesity as well as the CSTF sum, it also did not classify as many nonobese subjects as obese. Logically

this led to its slightly higher PPV (0.30).

The remaining two indicators, relative weight and the triceps skinfold, did vary by gender. The triceps skinfold indicator was less sensitive (79%) overall than the CSTF sum or relative BMI measures, but this reduced sensitivity was primarily a function of its use with females (64%), males being 93%. This reduced sensitivity in females, coupled with a specificity of 89%, resulted in a relatively unacceptable PPV of 0.24. In males, the PPV of the triceps skinfold was higher at 0.34.

The relative weight measure was found to be the least sensitive indicator overall (76%), with reduced sensitivity particularly evident with its use in males (67%). However, although relative weight did not identify "true" obesity as well as the other indicators, it did not classify as many non-obese subjects as obese, as evidenced by the slightly higher specificities (female-93% and male-91%), overall accuracies (females-93%; males-89%), and PPVs (females-0.40 and males-0.30).

Principal Components Solutions

A principal components analysis of 5 adiposity indicators (triceps, CSTF sum, relative weight, relative BMI, PI) performed on the total sample resulted in a unifactorial solution accounting for 79.4% of the variance (Table III.7). Factor loadings for each adiposity indicator ranged from 0.82 to 0.94. Principal component analyses for these same five indicators were repeated on subsamples of each gender, and on subsamples of subjects > and ≤ 11 years of age (Table III.7). All solutions were unifactorial, with percent of variance accounted for ranging between 78.4% and 85.7%, providing evidence that only one main

construct underlay all of the original five indicators of adiposity.

The criterion measure of body density for each subject was correlated with the factor scores derived from each of the above principal components solutions and with each of the original 5 adiposity indicators (Table III.8). For all solutions, the correlations with density were statistically significant, ranging from 0.43 to 0.81. By visual inspection a general pattern was evident wherein the CSTF sum and the triceps skinfold appeared to exhibit consistently stronger linear relationships with the criterion measure than with the factor scores. Statistical tests for differences in correlations between dependent groups confirmed that the CSTF sum was statistically more linearly related with the criterion measure than the factor scores in the following principal components solutions: total sample; sample > 11 years; males > 11 years; and females > 11 years. In the remaining 5 solutions, there were no such statistically significant differences in the linear relationship between the CSTF sum and factor scores with the criterion. In all solutions, the factor scores were more linearly related with the criterion measure than the height-weight indicators. It was reasonable to assume that if the factor solution was a true measure of adiposity the following condition should exist: factor scores derived from shared variance of all convenient indicators should correlate with the gold standard of adiposity, body density, more highly than any one of the convenient indicators which had entered the factor solution. Since this was only true for height-weight indicators and not skinfold indicators validation of the underlying construct as adiposity was therefore not evident in some cases.

Discussion

These results confirmed previous findings that the various convenient obesity indicators do not consistently identify subjects with obesity; others have suggested, depending on the definition and indicator of obesity used, that prevalence rates in children and adolescents vary between 3 and 25% (6,12,13,14). The proportions identified as obese in this study were very similar to our previous work (20), where between 4.9 and 18.8% of the sample were classified as obese, depending on the indicator used. Differences were present within height-weight indices, within skinfold indices, and among all of these and the criterion measure of percent body fat. Johnston (11) also noted inconsistencies between a percent body fat criterion measure and the triceps skinfold and relative weight indicators. However, as stated before (20), the presence of inconsistencies between indicators is not surprising since recommended obesity classification points in these various indicators have not been based on a common percentile point. This is an obvious requirement if reliable and valid classification of obesity is to be eventually derived.

It is interesting to note that the criterion measures used in our previous work and this study classified a low percentage of subjects as obese. In the previous study, 8.0% of the sample were "operationally" identified as obese, while in this study 5.4% were obese by the criterion percent body fat measure. Unfortunately in these two samples the visual rating scale and percent body fat measures were not used concurrently, therefore the apparent similarity in proportions identified as obese relate to independent samples and as such do not necessarily confirm the utility of visual rating as a criterion.

It was evident following examination of the diagnostic utility for obesity of each indicator that each should be used with some caution in identifying children and youths as obese. Although it is important to identify obese children for treatment and prevention of subsequent health risk, individuals should be aware of the clinical and psycho-social costs associated with each of these indicators, particularly when used to identify candidates for treatment programs. Therapy is frequently ineffective and relapse rate is high when there is initial success (35).

Two indicators (CSTF sum and relative BMI) did not vary by gender. Both indicators possess high sensitivities and therefore are good for identifying true obesity; obviously, if this is the primary purpose, these are the measures of choice. However, both have associated errors, particularly those related to psycho-social costs due to the inappropriate labelling of some non-obese children as obese. The low PPV of both tests (0.26, 0.30) indicates there is less than a 1 in 3 chance that the individual identified as obese will be truly obese.

The utility of the other indicators, triceps skinfold and relative weight, varies by gender. For males, the triceps skinfold does not identify true obesity as well as the CSTF sum, but its use is associated with slightly less psycho-social costs, as evidenced by the slightly higher specificity, accuracy, and PPV. The relative weight measure in males possesses a low sensitivity, and therefore its ability to detect true obesity is a concern, particularly in combination with its low PPV (0.30).

In females, the relative weight measure does not identify true obesity as well as the CSTF sum or relative BMI measures, but its use

is associated with lower psycho-social costs, as evidenced by its increased specificity and accuracy. Clinically, it appears to be one of the better indicators, with a 0.40 probability that a child identified as obese will be obese. Conversely, the triceps skinfold appears to be inappropriate for use in females: it does not identify true obesity well, and when a female is identified as obese, there is only a 1 in 4 chance that she will be obese. Therefore, prevalence rates for obesity that have been based on the use of the triceps skinfold in females could be 25% of what they are currently reported to be.

These ratings of the diagnostic utilities in this study, using percent fat as the criterion measure for obesity, differ from our previous work (20) in which a criterion measure was developed using a combination of factor scores and visual ratings. In that study, the positive predictive values were much higher, ranging from 0.43 to 1.00; the sensitivity for relative weight, although similar for the overall sample, was reversed for each gender, with females having reduced sensitivity, rather than males; the utility of the CSTF sum varied by gender, being more appropriate for use in females; and the relative BMI indicator had generally higher values for all diagnostic utilities. When a subsample consisting of those subjects greater than 11 y was analyzed to allow for a closer comparison to our previous work, many of the inconsistencies between the two studies remained. Johnston (11) has also compared relative weight and triceps skinfold measures to the same criterion of percent body fat used in the present study. He found that both were low in sensitivity (23-50%) and high in specificity (85-100%). Both indicators in this study have reduced sensitivity, but not to the extent that Johnston has reported.

Our two studies have produced conflicting evidence with regards to the diagnostic utility of the convenient indicators; the most likely explanatory hypothesis for this is that there are inherent differences in the use of visual inspection and percent body fat as criteria. The result of these differences are: 1. varying patterns in the diagnostic utilities of convenient indicators and thus different recommendations for their use between our previous work and this study; and 2. different proportions in each sample classified as obese (8.0 versus 5.4%). To determine if a comparable classification point to our previous work could be derived for body density and thus resolve the inconsistencies between the two studies, principal components analyses were performed.

The principal components analyses consistently indicated a unidimensional solution. This confirms our previous results (20) and suggests that a common construct underlies all 5 convenient indicators. Since all indicators have been used as measures of body composition, and our previous work suggested that the underlying construct was adiposity, it appeared reasonable to once again assume that the construct in this study was adiposity. However, in pursuing validation of this construct with body density as the criterion measure, it was demonstrated that an individual indicator (ie., CSTF sum) was statistically more linearly related to the criterion than the factor scores in 4 solutions, an inconsistent result for validating the underlying construct as adiposity and a problematic result for deriving an operational standard for obesity based on body density and factor scores.

Densitometry has been generally considered to be the "gold"

standard against which all new body composition methods should be judged (36,37,38); therefore, its use as a validating criterion measure of adiposity in this study was deemed appropriate. However, given the inconsistencies with the previous work and that the measurement of body density is difficult with children, questions regarding the accuracy of the body density measure remain. The main source of variability in body density in growing children is the biological variability of the fat free body mass, resulting from changes in water, mineral, and protein content (2). It was obvious that pubertal status varied greatly between subjects in this study. When genders were analyzed separately with the potential periods of rapid growth eliminated from the sample (females - aged 10 to 12 y; males - 12 to 14 y), the results suggested that the underlying construct was adiposity. That is, the factor scores were as or more linearly related to the criterion measure than any of the individual indicators that entered the solution. To confirm these results, separate analyses of females 10 to 12 y and males 12 to 14 y were conducted, with the expectation that the skinfold indicators would be more linearly related to the criterion than the factor scores and the other convenient indicators, as it was hypothesized that there would be more biological variability in the fat free body mass during these ages than others and hence greater variability in the body density measure. In addition, it is known that height-weight indicators are unreliable during periods of rapid growth (39); thus, factor scores which are more representative of height-weight indicators (3 height-weight indices versus 2 skinfold indices in the factor solution) are also likely to be less stable during such periods. The results of these analyses were as expected: during periods of rapid growth in both

males and females the CSTF sum of skinfolds was more linearly related to the criterion than the factor scores. Therefore, it appears that periods of rapid growth were a potential confounding influence in the validation of the underlying construct as adiposity. It is interesting to note that the S.E. for body density was largest in the identified potential periods of rapid growth for both males and females (.003 to .004 g/cc).

Thus, all factor scores were significantly correlated with the adiposity criterion, and in the majority of cases, statistical evidence suggested that the factor scores were at least as valid as measures of adiposity as any of the individual indicators.

However, the visual inspection rating scale used in our previous work appeared to discriminate the factor scores formed from the common variance of all measures better than the "accepted" criterion method of body density. Because individual indicators were more linearly related to the criterion than the factor scores in this study, it was inappropriate to use the factor scores in the development of an "operational" standard, as was done in the previous work. A study is required that involves the use of both density and visual inspection (to determine the differences between the two measures) and an assessment of pubertal status (to determine the appropriate indicator for use with pre-pubescent, pubescent, and post-pubescent subjects). Until performed, it is suggested that: 1. convenient indicators are associated with error; 2. there appears to be potential for inappropriate labelling and stigmatizing with all measures; and 3. a common percentile point for obesity should be agreed upon for all measures, so that comparable, reliable prevalence statistics can be

developed, and this common percentile can only be derived if an unquestioned criterion measure of adiposity is identified.

Table III.1: Age- and Gender-Specific Equations to Convert Body Density to Percent Body Fat (2)

Age	Female	Male
7 - 9	%F=(5.43/D - 5.03)	%F=(5.38/D - 4.97)
10 - 12	%F=(5.30/D - 4.89)	%F=(5.27/D - 4.85)
13 - 15	%F=(5.12/D - 4.69)	%F=(5.08/D - 4.64)

Table III.2: Age and Gender Distribution of the Sample

Age (yrs)	Gender		
	Female N	Male N	Total N
7	34	29	63
8	25	27	52
9	29	29	58
10	33	28	61
11	29	29	58
12	32	36	68
13	44	42	86
14	48	46	94
Total	274	266	540

Table III.3: Summary of Available Measures

Variable	Females All	Females > 11 y	Females ≤ 11 y	Males All	Males > 11 y	Males ≤ 11 y
	$\bar{x}$ ± SD (N = 274)	$\bar{x}$ ± SD (N = 124)	$\bar{x}$ ± SD (N = 150)	$\bar{x}$ ± SD (N = 266)	$\bar{x}$ ± SD (N = 124)	$\bar{x}$ ± SD (N = 142)
Triceps (mm)	14.3 ± 6.3	16.5 ± 6.6	12.4 ± 5.5	11.0 ± 4.9	11.5 ± 5.4	10.5 ± 4.3
Relative Weight (%)	102.1 ± 15.6	99.5 ± 16.3	104.3 ± 14.8	104.0 ± 13.0	103.7 ± 13.2	104.4 ± 12.9
Relative BMI (%)	105.4 ± 15.5	104.5 ± 15.4	106.2 ± 15.5	106.7 ± 13.8	107.0 ± 14.4	106.5 ± 13.4
BMI	18.4 ± 3.1	19.7 ± 2.9	17.3 ± 2.8	18.4 ± 2.8	19.8 ± 2.8	17.3 ± 2.3
PI	13.1 ± 0.6	13.2 ± 0.6	13.0 ± 0.6	13.1 ± 0.5	13.2 ± 0.5	13.0 ± 0.5
CSTF Sum (mm)	56.8 ± 23.9	61.9 ± 22.9	52.7 ± 23.9	46.2 ± 22.4	50.0 ± 24.5	42.9 ± 19.8
Density (g/cc)	1.0497 ± .015	1.0504 ± .015	1.0492 ± .016	1.0601 ± .016	1.0644 ± .016	1.0563 ± .015
% Body Fat	16.6 ± 7.5	17.8 ± 6.9	15.6 ± 7.9	12.8 ± 7.2	12.5 ± 6.9	13.1 ± 7.5

Table III.4: Frequencies and Percentages of the Sample Classified as Obese by the Obesity Indicators

INDICATOR	FEMALES		MALES		TOTAL SAMPLE	
	(N)	(%)	(N)	(%)	(N)	(%)
Relative Weight	30	10.9	33	12.4	63	11.7
Relative BMI	44	16.1	47	17.7	91	16.9
CSTF Sum	53	19.3	59	22.2	112	20.7
Triceps skinfold	37	13.5	41	15.4	78	14.4
% body fat	14	5.1	15	5.6	29	5.4

Table III.5: Z Values for Tests of Differences Between Proportions Classified as Obese and Nonobese

	Relative weight	Relative BMI	CSTF	Triceps	% body fat
Relative Weight					
Relative BMI	4.7*				
CSTF Sum	-5.7*	2.8*			
Triceps Skinfold	-1.9	1.8	4.7*		
% body fat	-4.9*	-7.6*	-9.1*	-6.3*	

* statistically significant at $@_2$=.05

Table III.6: Diagnostic Utilities of Four Adiposity Indicators in Identifying Obesity

Diagnostic Utility	Relative Weight	CSTF Sum	Relative BMI	Triceps Skinfold
A. Sensitivity				
Overall	76%	100%	93%	79%
Females	86%	100%	93%	64%
Males	67%	100%	93%	93%
B. Specificity				
Overall	92%	84%	87%	89%
Females	93%	85%	87%	89%
Males	91%	82%	87%	89%
C. Accuracy				
Overall	91%	85%	87%	89%
Females	93%	86%	88%	88%
Males	89%	83%	87%	89%
D. Positive Predictive Value				
Overall	.35	.26	.30	.29
Females	.40	.26	.30	.24
Males	.30	.25	.30	.34
E. Negative Predictive Value				
Overall	.99	1.00	.99	.99
Females	.99	1.00	.99	.98
Males	.98	1.00	.99	.99

Table III.7: Principal Components Solutions for Total Sample and Subsamples

Indicators	Solution 1 Sample	Solution 2 All Males	Solution 3 All Females	Solution 4 All>11 y	Solution 5 Males > 11	Solution 6 Females>11	Solution 7 All≤11 y	Solution 8 Males ≤ 11	Solution 9 Females≤11
Triceps	.82	.86	.83	.80	.84	.87	.87	.89	.87
CSTF Sum	.86	.89	.87	.86	.89	.89	.88	.91	.89
Relative Weight	.91	.95	.90	.88	.94	.87	.95	.96	.94
Relative BMI	.94	.95	.95	.93	.93	.95	.95	.96	.96
PI	.92	.91	.93	.95	.94	.95	.92	.91	.93
% of Variance	79.4%	83.0%	79.8%	78.4%	82.7%	81.8%	83.8%	85.7%	84.2%
N	540	266	274	248	124	124	292	142	150

Table III.8: Correlations Between Adiposity Indicators and Body Density Criterion for each Principal Components Solution

Indicator	Solution 1	Solution 2	Solution 3	Solution 4	Solution 5	Solution 6	Solution 7	Solution 8	Solution 9
	Sample	All Males	All Females	All>11 y	Males > 11	Females>11	All≤11 y	Males ≤ 11	Females≤11
Triceps	.71	.64	.71	.77	.72	.73	.74	.64	.81
CSTF Sum	.73	.66	.77	.81	.79	.80	.75	.68	.80
Relative Weight	.53	.57	.60	.43	.55	.54	.62	.61	.66
Relative BMI	.56	.53	.69	.48	.48	.67	.64	.62	.70
PI	.57	.56	.63	.54	.56	.61	.58	.53	.66
Factor Scores	.69	.65	.76	.68	.68	.74	.73	.66	.79
N	540	266	274	248	124	124	292	142	150

Bibliography

1. Burton BT, Foster WR, Hirsch J, Van Itallie TB. Health implications of obesity: an NIH Consensus Development Conference. Int J Obesity 1985;9:155-169.

2. Lohman TG. Applicability of body composition techniques and constants for children and youths. In: Pandolf K, eds. Exercise and Sports Sciences Reviews 1986;14:325-357.

3. Wilmore JH, Buskirk ER, Digirolamo M, Lohman TG. Body composition. A round table. Phys & Sportsmed 1986;14(3):144-162.

4. Manson JE, Stampfer MJ, Hennekens CH, Willett WC. Body weight and longevity. A reassessment. JAMA 1987;257(3):353-358.

5. Simopoulos AP. The health implications of overweight and obesity. Nutr Reviews 1985;43(2):33-40.

6. Dietz WH. Childhood obesity: susceptibility, cause, and management. J Pediatrics 1983;103:676-685.

7. Paige DM. Obesity in childhood and adolescence. Special problems in diagnosis and treatment. Postgrad Med 1986;79(1):233-245.

8. Garrow J. Obesity - Public health enemy No. 2? Health Educ J 1986;45(1):57-58.

9. Roche A, Siervogel R, Chumlea W, Webb P. Grading body fatness from limited anthropometric data. Am J Clin Nutr 1981;34:2831-2838.

10. Rolland-Cachera MF, Sempe M, Guilloud-Bataille M, Patois E, Pequinot-Guggenbuhl F, Fautraud V. Adiposity indices in children. Am J Clin Nutr 1982;36:178-184.

11. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. J Adolesc Health Care 1985;6(3):185-190.

12. Hager A. Nutritional problems in adolescence - obesity. Nutr Rev 1981;39(2):89-95.

13. Franklin B, Rubenfire M. Losing weight through exercise. JAMA 1980;244(4):377-379.

14. Heston M. Childhood obesity: Implications for physical education. Runner 1985;23(3):3-6.

15. Moynihan CM, Stark O, Peckham CS. Obesity in 16-year-olds assessed by relative weight and Doctors' rating. Int J Obesity 1986;10:3-10.

16. Stark O, Lloyd JK. Some aspects of obesity in childhood. Postgrad Med Journal 1986;62:87-92.

17. Brownell K, Kaye F. A school-based behaviour modification, nutrition education, and physical activity program for obese children. Am J Clin Nutr 1982;35:277-283.

18. Bray GA. Overweight is risking fate. Definition, classification, prevalence, and risks. Ann NY Aca Sci 1987;499:14-28.

19. Van Itallie TB. Health implications of overweight and obesity in the United States. Ann Intern Med 1985;103(6 pt 2):983-988.

20. Marshall JD, Hazlett CB, Spady DW, Quinney HA. A comparison of convenient indicators of obesity. Am J Clin Nutr In Press.

21. Poskitt E. Management of obesity. Arch Dis Child 1987;62:305-310.

22. Canadian Standardized Test of Fitness (CSTF) Operations Manual. Ottawa: Minister of State, Fitness and Amateur Sport, 1986.

23. Ross W, Ward R. The O-scale system. Vancouver:Rosscraft Ltd., 1984.

24. Canada Fitness Survey: Canadian Youth Fitness. Ottawa:Fitness Canada, Fitness and Amateur Sport, 1985.

25. Johnston JL, Leong MS, Checkland EG, Zuberbuhler PC, Conger PR, Quinney HA. Body fat assessed from body density and estimated from skinfold thickness in normal children and children with cystic fibrosis. Am J Clin Nutr 1988;48:1362-1366.

26. Katch F, Micheal E, Horworth S. Estimation of body volume by underwater weighing: description of a simple method. J Appl Physiol 1967;28:811-815.

27. Bass B. Lung Function Tests. London:HR Lewis & Co., Ltd., 1974.

28. Meneely G, Ball C, Kory R, et al. A simplified closed circuit helium dilution method for the determination of the residual volume of the lungs. Am J Med 1960;28:824-831.

29. Hamill P, Drizd T, Johnson C, Reed R, Roche A, Moore W. Physical growth: National center for health statistics percentiles. Am J Clin Nutr 1979;32:607-629.

30. Micozzi MS, Albanes D, Jones DY, Chumlea WC. Correlations of body mass indices with weight, stature, and body composition in men and women in NHANES I and II. Am J Clin Nutr 1986;44:725-731.

31. Gravetter FJ, Wallnau LB. Statistics for the Behavioural Sciences. New York:West Publishing Company, 1988.

32. Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology - the essentials. London: Williams & Wilkins, 1988.

33. Jackson AS. Research design and analysis of data procedures for predicting body density. Med Sci Sport Exer 1984;16(6):616-620.

34. Harman HH. Modern Factor Analysis. Chicago: The University of Chicago Press, 1967.

35. Brownell KD. The psychology and physiology of obesity: Implications for screening and treatment. J Am Dietit Assoc 1984;84(4):406-414.

36. Brodie DA. Techniques of measurement of body composition. Part I. Sports Med 1988;5:11-40.

37. Oppliger RA, Looney MA, Tipton CM. Reliability of hydrostatic weighing and skinfold measurements of body composition using a generalizability study. Hum Biol 1987;59(1):77-96.

38. Roche AF. Some aspects of the criterion methods for the measurement of body composition. Hum Biol 1987;59(2):209-220.

39. Johnston FE. Relationships between body composition and anthropometry. Hum Biol 1982;54(2):221-245.

Chapter IV: General Discussion and Conclusions

Obesity has been defined as an excess of body fat leading to a significant impairment of health (1). However, the measurement of fat in the human is difficult, and the point (or range) at which fat becomes clinically hazardous is a subject of considerable debate (2,3). This is particularly true in children, where pathological sequalae of fatness occur infrequently (4). Therefore, definitions of obesity are statistically based and somewhat arbitrarily derived.

The need for a quantifiable definition of obesity is great, and its availability basic to the accurate establishment of prevalence statistics and the organization of obesity treatment programs. Obviously the ideal definition would involve an identification of a threshold amount of fat that confers excess health risk. The types of studies required to derive this definition have not been done. However, as Dietz and Gortmaker (4) state:

> ". . .The statistical definition of a disease based on a continually distributed variable such as fatness is not unique to obesity. The definition of anemia, hyperlipidemia, and hypertension are also dichotomous, statistically derived, and somewhat arbitrary."(p. 122).

Obesity may also be related to other factors affecting the quality of life (i.e., physical activity, social well-being, mental health) and therefore obesity cut-off points may vary in populations and throughout the life cycle. Investigators have demonstrated that obesity in adults is related to increased mortality and morbidity (5,6), but others suggest that obese people are not always unhealthy (4,7). Health risks of obesity are poorly established in children, but obesity in childhood predisposes to obesity in adulthood, where the risks are more clear (8,9,10). Therefore, statistically and somewhat

arbitrarily derived definitions of obesity will identify a population at potential increased health risk (4), but will also include some persons who are overweight and without medical risk (ie., false positives).

Errors (ie., false positives, false negatives) will occur with the use of convenient obesity indicators. Awareness of the particular type of errors associated with each indicator is particularly important when the indicator is used to identify candidates for treatment programs. Therapy is frequently ineffective and relapse rate is high when there is initial success (2,3). There are a number of potential hazards associated with obesity treatment, some of which include: the development of eating disorders or restrained eating patterns; unrealistic expectations and the development of an unrealistic body image; the promotion of the idea that self-worth and self-acceptance are dependent on attaining "ideal" body size and shape (11); and metabolic problems associated with unusual diet patterns. It would be optimal, therefore, if candidates selected for obesity treatment programs were truly at risk, since the obesity treatment effects in some cases may be more damaging than the obesity itself (12). Currently, it would appear individuals who are at risk are those: 1. who are grossly obese, and/or 2. who have "centralized" or an android-type of obesity (5), although this latter point has yet to be confirmed in adolescents.

Commonly used convenient indicators provide, for the most part, an indication of the degree of adiposity, rather than an indication of a particular fatness distribution pattern. A number of these have

accepted obesity classification points, while others have more than one (2,3,9). The classification points are not based on a common percentile point, and therefore indicators classify the same individual differently, resulting in confusing prevalence statistics and problems regarding admission to and evaluation of obesity treatment programs. Similar classification criteria are often used for children, adults, and for both genders, although there is little justification for this (13). Validation work with these convenient indicators has been insufficient (14,15,16).

It was in reference to these problems that two studies were undertaken to examine the validity of convenient obesity measures. In the first study a "gold standard" measurement for obesity was unavailable and consequently an "operational standard" was derived using the common variance of all studied convenient indicators (i.e., factor scores) in conjunction with a visual rating. This operational standard was used to determine the diagnostic utility of four convenient obesity indicators: relative weight (17), relative BMI (18), the CSTF sum (19), and the O-scale (20). The results suggested that the utility of the four indicators varied and that the use of most should not be considered independent of the gender of the subject, or without reference to the purpose for classifying subjects as obese. In the second study, a "gold standard" measurement was available: percent body fat estimated from underwater weighing (9,16,21). The utility of four convenient indicators, relative weight, relative BMI, the CSTF sum, and the triceps skinfold, were derived using the dichotomized percent body fat measure (in reference to

obesity) as a gold standard. Unfortunately, the visual rating scale and percent body fat measures were not used concurrently in the second study, as a retrospective data set was used, making direct comparisons between the two studies difficult. However, similar statements regarding prevalence rates and the underlying construct of convenient obesity indicators can be made from both studies.

Prevalence of Obesity

Investigators have suggested that the prevalence of childhood and adolescent obesity has increased drastically over the last twenty years (22,23). Dietz and Gortmaker (4), defining obesity as a triceps skinfold greater than or equal to the 85th percentile for children and adolescents of the same age and gender studied during the National Health Examination Surveys Cycles 2 and 3, demonstrated that the prevalence of obesity in youngsters in the United States increased by over 40% between 1965 and 1980. The trend of increasing obesity prevalence in the United States does not appear to be the case over the past 9 years in Edmonton children. Based on this research the prevalence of obesity in Junior High School children has remained relatively constant, or slightly decreased, over the past 9 years (Table IV.1).

Obesity has also been described as reaching epidemic proportions in some countries (22,24). This research does not support this contention in Canada. If the two criterion measures from both studies, factor scores and percent body fat, are used to describe prevalence, only 8.0% and 5.4% of the samples studied were obese. This is far from the 25 to 40% that others have reported (2,24).

However, support is provided for the idea that different indicators classify obesity differently. With few exceptions in both studies, there was a statistically significant difference in the proportion of the samples classified as obese between all indicators. This may have resulted because the indicators measure different aspects of body composition (eg., height and weight versus skinfolds) or because the obesity classification point for each indicator was not based on a common percentile point. It was demonstrated in both studies, through the use of principal components analyses, that one main construct underlay all of the indicators, and that this construct logically and statistically appeared to be adiposity. Therefore, it appeared that indicators classify obesity differently because obesity classification points in the various indicators have not been based on a common percentile point. Without this, confusing obesity prevalence rates abound. In addition, many children and adolescents may be subjected to unnecessary treatment and resulting psychological trauma because of inaccurate diagnosis. The identification of a common percentile point for the classification of obesity for various convenient indicators is necessary to ensure reliability in classification, and as a prerequisite to a valid classification of obesity.

Factor Solutions

Despite the fact that slightly different measurements were examined in the first versus the second study, principal components analyses consistently demonstrated a unifactorial solution, accounting for approximately 80% of the variance, providing evidence that only one main construct, adiposity, underlay all of the original indicators studied. To verify this, validation work was pursued with selected

criterion measures of adiposity: visual rating and body density in the first and second studies, respectively. The presence of the following condition was necessary for the factor solution to be deemed a true measure of adiposity: factor scores derived from the shared variance of all convenient indicators should have been more highly correlated with the criterion measures than any one of the convenient indicators which had entered the factor solution. In the first study, this condition was met, with the factor scores having the strongest linear relationship (0.81) with the criterion visual rating than any of the individual indicators. In the second study, validation of the underlying construct was not as evident; in some solutions the skinfold measures were more highly correlated with the criterion measure than the factor scores. However, there was an unequal distribution of measurement methods in the second study (3 height-weight indices, 2 skinfold indices) and thus factor scores were more representative of height-weight indices than skinfold indices. In addition, puberty appeared to be a confounding influence. When periods of rapid growth (females 10-12 y; males 12-14 y) were accounted for, the results suggested that the underlying construct was adiposity. Therefore, in both studies, there was statistical evidence to support the hypothesis that the underlying construct of the studied convenient indicators was adiposity.

Both skinfold indices and height-weight indices are considered to be "doubly indirect" convenient measures of adiposity. However, investigators have suggested that skinfolds are a more "direct" way of measuring adiposity, as they are a measure of fatness per se and height-weight indicators may be compounded by other variables such as muscle mass or skeletal size (24). This was supported by this research:

in both studies the CSTF skinfold sum was more highly related to the respective criterion measure of adiposity than the height-weight indicators. A sum of skinfolds also appears to be more representative of fatness than a single skinfold: the CSTF sum was more highly correlated with the criterion measure of adiposity than the triceps skinfold in both studies. This supports the research of others (25,26,27,28).

Diagnostic Utilities

Since the studied convenient indicators were based on a common underlying construct, adiposity, their use as convenient measures of obesity, which involves a measure of adiposity, was justified. Indeed, many convenient measures of adiposity have obesity classification points. However, it was apparent, through an examination of the diagnostic utilities of each indicator in comparison to an "operational" or "gold" standard for obesity that: 1. the classification points for each indicator were different; and 2. the use of each measure should not be considered independent of the gender of the subject or without reference to the purpose for classifying individuals as obese.

The use of different criterion measures in each study resulted in inconsistencies in the diagnostic utilities of each convenient indicator. The lack of a clearly accepted criterion measure for obesity is the key methodological shortcoming in the validation of convenient indicators. Obviously, a study involving the use of both densitometry and visual rating is required to allow for a direct comparison between the use of the two criterion measures.

Percent body fat was chosen as the "gold" standard for obesity

classification for this research; although admittedly imperfect, its use was supported by the work of others (9,16,21,29). However, it is unlikely that the use of a single cut-off point for obesity (eg., males-25%; females-30%) is justified across the age ranges studied. This is supported by data from the second study (Table IV.2). For example, if an individual were to statistically and arbitrarily define obesity as the upper 15% of the population, as is commonly done, the percent fat cut-off point in females would vary between 20.5% (7-8 yr olds) and 28.1% (11-12 yr olds). In other words, 25% body fat may represent obesity in a 7-10 year old girl, but normal growth in an 11-12 year old. Therefore, it appears that the use of a single cut-off point as a "gold" standard is not justified and that obesity cut-off points should be age-, gender-, and population-specific. The use of a single cut-off point across age ranges will affect the diagnostic utilities of the studied indicators. As Fletcher et al (30) state:

> ". . .if an inaccurate standard of validity is used, a new test can perform no better than that standard and will seem inferior when it approximates the truth more closely."(p. 46)

There were inconsistencies in the sensitivity and specificity of each indicator between the first and second studies, making the choice of the "best" indicator difficult. However, the selection of a sensitive or specific test depends on the perspective of the user. For example, sensitive tests are selected when there is an important penalty for missing a disease (30). An individual who believes that obesity results in increased health risk and therefore feels early detection and treatment is important would be best to choose a sensitive test. A sensitive test is most helpful when the test result is negative, i.e., "rules out" disease (30). Conversely, specific tests

are useful to confirm a diagnosis suggested by other data. A specific test rarely gives false positive results, and therefore is useful when false positive results can harm patients physically, emotionally, or financially (30). Individuals who come from more of a psycho-social perspective and who believe that there are hazards associated with obesity treatment would be better to choose a specific test. There is a trade-off between the sensitivity and specificity of a test, particularly when the location of a cut-off point is an arbitrary decision (30), as is the case with obesity. It is important to note that the sensitivities and specificities reported for the studied indicators are only for the subjects and age ranges examined; the diagnostic utility of these indicators may vary with use in other samples of different ages and background.

Predictive values represent post-test probabilities and are determined by the sensitivity and specificity of the test and the prevalence of disease or condition in the population being tested. Sensitive tests result in higher negative predictive values, while specific tests result in higher positive predictive values (30). However, because predictive value is also influenced by prevalence, it is not independent of the setting in which the test is used. As Fletcher et al (30) state:

> ". . as the prevalence of disease in a population approaches zero, the positive predictive value of a test also approaches zero. Conversely, if everyone in a population tested has the disease, all negative results will be false negatives, even for a very sensitive test. As prevalence approaches 100%, negative predictive value approaches zero." (p. 56).

Since the two criterion measures used in the first and second study resulted in slightly different obesity prevalence rates (8.0 and 5.4%, respectively), different positive predictive values between the two

studies were not unexpected. A lower prevalence rate was found in the second study, consequently the positive predictive values were lower than those reported in the first study. Both studies reported high negative predictive values for all indicators; this is not unexpected as the prevalence rates for both studies were relatively low.

Concluding Remarks

Convenient obesity indicators appear to measure adiposity. However, because obesity cut-off points for the various indicators have been arbitrarily determined and have not been based on a common percentile point, indicators possess different diagnostic utilities.

The main methodological shortcoming in the validation of convenient obesity indicators is the lack of availability of a "gold standard" measure of obesity. Until a "gold standard" is identified and generally accepted, the true validity of convenient obesity indicators is difficult to establish. Therefore, one of the main thrusts of research in this area should be the identification of a "gold standard" measure of obesity that is known to confer excess health risk. Once identified, the validity of convenient indicators should be established, and the identification of a percentile point for obesity classification be agreed upon for all measures, so that comparable, reliable prevalence statistics can be developed. Until such time, researchers should be aware of the inherent errors associated with the use of convenient measures and the potential for inappropriate labelling and stigmatizing. These errors, as identified in both studies of this thesis, should be established in both genders at various age groups. Cut-off points for obesity should be identified separately for each gender. Additionally, obesity cut-off points may vary with the type of

risk being examined. For example, physical activity risk (i.e., decreased physical activity due to fat content) may occur at a lower fat content than some of the more traditional medical risks, and therefore investigators should be aware of the type of risk they are examining.

It also appears that puberty is a confounding influence in the use of convenient indicators. Further research with this age group should include a measure of pubertal status, and appropriate indicators should be identified for use with pre-pubescent, pubescent, and post-pubescent subjects.

Finally, convenient obesity indicators are often used as screening tools for obesity treatment programs, and since each is associated with inherent error, the use of a combination of indicators (eg., relative BMI with skinfolds; relative BMI with visual rating) may be more effective in the identification of appropriate candidates for treatment than the use of a single indicator. A full pre-treatment screening should include these types of measures in addition to measures of motivation, social support, self-esteem, body image, physical activity, dietary history, and family history.

Table IV.1: Proportion of Samples Classified as Obese by Various Indicators

Indicator	1981 Data (Sample > 11 y)		1988 Data	
	FEMALES	MALES	FEMALES	MALES
CSTF Total	21.0%	19.4%	15.9%	21.5%
Relative Weight	8.9%	14.5%	9.9%	10.8%
Relative BMI	15.3%	19.4%	16.2%	13.7%

Table IV.2: Percent Body Fat Cut-Off Points at Various Percentiles in Males and Females (Study 2).

Age Ranges	N	Percentiles				
		50	80	85	90	95
Females						
7 - 8	59	12.6	17.4	20.5	23.0	24.5
9 - 10	62	15.9	22.1	22.9	27.5	31.1
11 - 12	61	15.8	26.6	28.1	31.3	34.9
13 - 14	92	18.4	23.0	24.4	26.8	29.3
Males						
7 - 8	56	11.1	18.6	19.9	21.9	24.4
9 - 10	57	11.0	18.9	20.4	22.3	26.0
11 - 12	65	13.3	20.7	22.4	24.3	31.2
13 - 14	88	11.2	17.5	19.4	23.6	26.1

Bibliography

1. Burton BT, Foster WR, Hirsch J, Van Itallie TB. Health implications of obesity: an NIH Consensus Development Conference. Int J Obesity 1985;9:155-169.

2. Dietz WH. Childhood obesity: susceptibility, cause, and management. J Pediatrics 1983;103:676-685.

3. Paige DM. Obesity in childhood and adolescence. Special problems in diagnosis and treatment. Postgrad Med 1986;79(1):233-245.

4. Dietz WH, Gortmaker SL. Reply to letter. AJDC 1988;142:121-122.

5. Manson JE, Stampfer MJ, Hennekens CH, Willett WC. Body weight and longevity. A reassessment. JAMA 1987;257(3):353-358.

6. Simopoulos AP. The health implications of overweight and obesity. Nutr Reviews 1985;43(2):33-40.

7. Callaway W. Weight standards: their clinical significance. Ann Intern Med 1984;100(2):296-298.

8. Brownell K, Kaye F. A school-based behaviour modification, nutrition education, and physical activity program for obese children. Am J Clin Nutr 1982;35:277-283.

9. Bray GA. Overweight is risking fate: Definition, classification, prevalence, and risks. Ann NY Aca Sci 1987;499:14-28.

10. Van Itallie TB. Health implications of overweight and obesity in the United States. Ann Intern Med 1985;103(6 pt 2):983-988.

11. Wooley SC, Wooley OW. Should obesity be treated at all? In: Stunkard AJ, Stellar E (eds.). Eating and Its Disorders. New York: Raven Press. 1984;185-192.

12. Brownell KD. The psychology and physiology of obesity: Implications for screening and treatment. J Am Dietit Assoc 1984;84(4):406-414.

13. Moynihan CM, Stark O, Peckham CS. Obesity in 16-year-olds assessed by relative weight and Doctors' rating. Int J Obesity 1986;10:3-10.

14. Roche A, Siervogel R, Chumlea W, Webb P. Grading body fatness from limited anthropometric data. Am J Clin Nutr 1981;34:2831-2838.

15. Rolland-Cachera MF, Sempe M, Guilloud-Bataille M, Patois E, Pequinot-Guggenbuhl F, Fautraud V. Adiposity indices in children. Am J Clin Nutr 1982;36:178-184.

16. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. J Adolesc Health Care 1985;6(3):185-190.

17. Schwartz M. Childhood obesity. Surgical Clin North Am 1979;59(6):995-1005.

18. Poskitt E. Management of obesity. Arch Dis Child 1987;62:305-310.

19. Canadian Standardized Test of Fitness (CSTF) Operations Manual. Ottawa: Minister of State, Fitness and Amateur Sport, 1986.

20. Ross W, Ward R. The O-scale system. Vancouver: Rosscraft Ltd, 1984.

21. Wilmore JH, Buskirk ER, Digirolamo M, Lohman TG. Body composition. A round table. Phys & Sportsmed 1986;14(3):144-162.

22. Gortmaker SL, Dietz WH, Sobol AM, et al. Increasing pediatric obesity in the United States. AJDC 1987;141:535-540.

23. Kashani I. Obesity in children. AJDC 1988;142:121.

24. Dietz WH. Obesity in infants, children, and adolescents in the United States. I. Identification, natural history, and aftereffects. Nutr Research 1981;1:117-137.

25. Pollock ML, Jackson AS. Research progress in validation of clinical methods of assessing body composition. Med Sci Sport Exer 1984;16(6):606-613.

26. Clarys JP, Martin AD, Drinkwater DT, Marfell-Jones MJ. The skinfold: myth and reality. J Sports Sciences. 1987;5:3-33.

27. Shear CL, Freedman DS, Burke GL, Harsha DW, Berenson GS. Body fat patterning and blood pressure in children and young adults. The Bogalusa heart study. Hypertension. 1987;9:236-244.

28. Wadden TA, Stunkard AJ, Johnston FE, et al. Body fat deposition in adult obese women. II. Changes in fat distribution accompanying weight reduction. Am J Clin Nutr 1988;47:229-234.

29. Brodie DA. Techniques of measurement of body composition. Part I. Sports Med 1988;5:11-40.

30. Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology - the essentials. London: Williams & Wilkins, 1988.

APPENDIX 1
REVIEW OF RELATED LITERATURE

REVIEW OF RELATED LITERATURE

TABLE OF CONTENTS

	PAGE
I. The Measurement of Body Fat	78
1. Laboratory Techniques	78
A. Two Compartment Model	79
B. Specific Techniques	81
i. Densitometry	81
ii. Total Body Water	88
iii. Total Body Potassium	89
iv. Other Laboratory Techniques	91
a. Muscle Mass Estimations	91
b. Nuclear Magnetic Resonance Imaging	93
c. Bioelectrical Impedance	94
d. Total Body Electrical Conductivity	95
e. Ultrasonics/Radiographics	96
f. Photon Absorptiometry	97
g. Infrared Interactance	98
h. Multicomponent Models	98
2. Clinically Convenient Indicators	99
A. Skinfolds	100
i. Prediction of Body Fat from Skinfolds	104
a. Population-Specific Approaches	105
b. Generalized Approaches	105
c. Estimation in Children	106
d. Error	107
e. Alternative Uses of Skinfolds	108
B. Height-Weight Indices	109
i. Height-Weight Charts	109
ii. Body Mass Index	111
iii. Ponderal Index	113
iv. Use of Height-Weight Indicators in Children	114
C. Visual Estimation	115
3. Summary - Body Composition Methodology	116
II. What Amount of Fat is Considered Excessive?	117
1. Statistical Definitions of Obesity	118
2. Comparative Studies	121
3. Measurement of Obesity in Children	123
4. Health Risks of Obesity	124
5. Health Risks of Obesity in Children	127
III. Growth and Obesity	127

1. Growth of Fat Tissue 128
2. Growth of Lean Body Mass 130
i. Water Content of FFB 130
ii. Mineral Content of FFB 131
iii. Potassium Content of FFB 131
iv. Protein Content of FFB 132

IV. Principal Components Analysis 132

V. References 140

Appendix 1:Review of Related Literature

Obesity, a disorder of energy imbalance that results from multifactorial causes, has been defined as "an excess of body fat frequently resulting in a significant impairment of health" (p. 157) (23). This definition leads to two important questions: 1. how is body fat measured?; and 2. what amount of fat is considered to be excessive? The primary purpose of this literature review is to determine the contemporary answers to these questions. The secondary purpose is to discuss obesity within a growth and development context.

I The Measurement of Body Fat

All methods of body composition measurement are considered to be indirect estimations for which accuracy is affected by measurement error and biological variability (17,18,108). Wilmore (180) suggested that there were 3 levels of expertise in body composition measurement: techniques used in clinical practice (e.g., relative weight); more advanced techniques used in sports medicine or health surveys (e.g., underwater weighing, bioelectrical impedence); and advanced research laboratory techniques (e.g., TOBEC, NMR). For the purpose of this review, body composition measurement techniques will be considered to be one of two types: 1. laboratory techniques, or 2. clinically convenient techniques.

1. Laboratory Techniques

Several indirect laboratory measures of adiposity have been developed, based on a variety of theoretical relationships and the use of "constants" developed through human and animal cadaver dissections (108). These methods include: densitometry; measurement of total body water; measurement of total body potassium; measurement of creatinine

and 3-methylhistidine excretion; ultrasound; computed tomography; nuclear magnetic resonance imaging; total body electrical conductivity; electrical impedence; and infrared interactance (17,18,112). While they may provide moderately precise to accurate estimates of body fat (compound error of 3-4%) (18,108), these methods are also expensive, clinically inconvenient or unavailable, time-consuming, and some require a high degree of operator training and potential subject discomfort. In addition, the "constants" that are used are not always generalizable to the population of interest (12,13,17,18,80,81,108, 111). A major limitation to many of these methods is the fact that they are based on the two compartment model of body composition.

A. Two Compartment Model

There are 2 different two compartment models of body composition, both of which suggest that the body is composed of two primary constituents, each of which is assumed to be of constant composition and density for all individuals in a given population (39,100,108,109, 112,158,180). Behnke's (9) two compartment concept consisted of a division of adipose tissue and lean body mass (LBM), whereas Siri's (158) and Brozek's (22) concept was a division of fat and fat-free body (FFB).

Many investigators use FFB and LBM interchangably, but this is not correct. The FFB is composed of the weight of all tissues in the body minus the extractable fat and is estimated, through human and animal cadaver analysis and the results of indirect estimates of FFB composition studies in humans, to be composed of 73.8% water, 19.4% protein, and 6.66% mineral, and to have a density of 1.1 g/cc (108,180). Conversely, LBM is larger than the FFB because it includes

essential lipids in combination with the FFB, and therefore has a density and water content somewhat less than the FFB (4,108,166,180). The estimate of fat in LBM is between 1.5-3.0% (4,180). The constants originally developed for use in body compostion studies are based on the chemical composition of the FFB, not LBM (24,108). Therefore, throughout this review, unless stated, estimations will refer to FFB, and not LBM. The fat compartment in the two compartment model is assumed to be anhydrous with a density of .9 g/cc.

The primary advantage of the two compartment model is that, through the use of developed constants and relationships, it allows for the measurement of only one constituent of the FFB from which the relative fat and FFB content can be estimated (109).

However, there are limitations associated with the use of this model, the major one being the assumption of constant composition and density of the FFB. The composition and density of the FFB is affected by disease states, obesity, activity levels (and therefore muscle and bone development), age, maturation level, and gender (12,13,17,18,29, 30,31,53,54,56,80,81,108,111,114,161,166,170,177,180,184). Therefore, as Lohman (108) has suggested, the assumption that the density of the FFB is constant may be violated more frequently than is commonly thought. Other limitations are that separate estimates of the constituents of FFB are not made, individual prediction errors of fat content can be substantial, and that estimates of body composition in populations other than young adult males may systematically underestimate or overestimate fatness (109). The variability of the FFB has been recognized in male children (12,13,29,54,80,81,108,111) and young male adults, but needs to be defined in females, athletes,

the elderly, the disabled, and for those individuals with different ethnic and racial backgrounds (108,144).

Therefore, one must be aware of these limitations when using body composition methods that rely on the assumptions of the two compartment model. It is estimated that these methods are associated with an error of approximately 3-4% in the prediction of percentage of body fat (108, 180) because the assumptions of the two compartment model are violated.

B. Specific Techniques

Three standard laboratory techniques used to estimate body composition include densitometry, total body water, and total body potassium. These techniques will be discussed in some detail, while a brief review of other, more recent techniques will be provided.

1. Densitometry

Densitometry is generally considered to be the "gold standard" against which all new body composition methods should be validated (17, 31,47,134,144,167,182). However, densitometry is not without error (108,181), and therefore should not be considered as the sole criterion method.

Body density (D) is calculated from the relationship between body weight (W) and volume (V):

$$D = W/V \qquad (1)$$

If the body is assumed to be composed of two constituents of different densities, as is done in the two compartment model, the density of the whole system can be determined by the proportions of the two components in the system (9). The density of a system which is comprised of constituents A and B with respective densities a and b is:

$$D = \frac{A + B}{(A/a) + (B/b)} \qquad (2)$$

If it is assumed that the whole system equals unity, the proportional contribution of B (B = 1 - A) would be(9):

$$B = \frac{1}{D} X \frac{ab}{(a - b)} - \frac{b}{(a - b)} \qquad (3)$$

If in equation 2, A represents the proportion of lean tissue and B represents fat, the relative and absolute fat and lean components can be calculated quite easily. For example, Brozek et al (21) and Siri (158) have used this approach to derive equations for estimating body fat from measurements of body density.

However, to estimate body density to allow for the subsequent calculation of body fat and lean body mass, is is necessary to obtain a measurement of body volume. Body volume can be measured by hydrostatic weighing, water displacement, or helium dilution (9).

Hydrostatic weighing is the most widely used method for assessing body volume (4,9,96,144). The technique utilizes Archimedes' prfnciple: a body immersed in a fluid is acted on by a bouyancy force, which is evidenced by a loss of weight equal to the weight of the displaced fluid (9). Therefore, when a subject is weighed underwater while totally submerged, their total body volume is equal to their loss of weight in water, corrected for the density of water (D_w) corresponding to the water temperature at the time of the underwater weighing (9). Volume (V) is described as:

$$V = \frac{W_a - W_w}{D_w} \qquad (4)$$

where W_a and W_w represent the subject's weight in air and water, respectively. Equation 1 then becomes:

$$D = \frac{W_a}{\frac{W_a - W_w}{D_w}} \qquad (5)$$

However, equation 4 includes two extraneous volumes (residual volume of lungs, intestinal gas) in the calculation of total body volume which should be accounted for in a revised estimate of total body volume. A correction must be made for air remaining in the lungs at the time of underwater weighing (commonly residual volume). The residual volume (RV) makes a sizable contribution in the estimate of the total body volume, and since it is highly variable, it is essential to obtain a close approximation of the subject's actual RV. A less important source of error is the amount of intestinal gas, and since it is impractical to measure, estimates of between 30 to 100 ml are often used (9, 144).

When RV (or appropriate lung volume) and intestinal gas are accounted for, equation 5 becomes (9):

$$D = \frac{W_a}{\frac{(W_a - W_w)}{D_w} - (RV + 100\ ml)} \qquad (6)$$

The underwater weighing technique is associated with 2 sources of error: technical (measurement) error and biological variability. There are four sources of technical error associated with underwater weighing: 1. weight in air; 2. weight in water; 3. water temperature; and 4. residual volume (144). Weight in air and the water temperature can be measured relatively accurately with the appropriate equipment. There may be large errors in the measurement of weight in water, but these can be reduced through the use of a load cell over an autopsy scale, since the analogue output can be used to make more precise measurements from the fluctuating readings (17). Subjects should be

several hours post-prandial, clean, with all air expelled from their garments, and have defecated and urinated immediately prior to being measured (17). The common procedure is to weigh the subject in water 10 times and employ the mean of the last 3 or 4 repetitions (134,144). Apparently, a learning effect is present during the first several repetitions which increases the measurement error for the 10 repetitions, but reliability is enhanced if only the last 4 repetitions are used (134). Carbonated drinks can increase this source of error, and therefore should not be drunk before testing (17).

The accurate estimation of total body volume by hydrostatic weighing depends on the accurate measurement of individual lung volumes (178). The fourth source of technical error is the error of measurement of lung volume. Most hydrostatic weighing procedures involve subjects weighed at residual volume (RV) (8,39,93,134,150,161,171) because it is thought to be the lung volume least affected by hydrostatic pressure. However, expiring to RV underwater or on land is difficult for some people (178) and therefore alternatives, such as functional residual capacity (FRC) (53) and total lung capacity (TLC), have been examined (60,122,178). Weltman et al (178) found that hydrostatic weighing at TLC in prepubertal males resulted in a significantly lower determination of total body volume as compared to hydrostatic weighing at RV, and therefore concluded that the use of hydrostatic weighing at TLC was unacceptable when applied to prepubertal males.

Lung volumes can be measured with the subject totally submerged (12,39,93,110,134,146,161), with the head out of the water (44), or on land with the subject in the same position as when weighed in water (8,12,36,89,150,161,171,178). There is controversy as to which method

is best (12,17,144). RV measured with the body submerged has been demonstrated to be increased, decreased, and remain the same as when measured on land (17). However, as Roche (144) states

> "It does not matter whether the "residual volumes" measured "on land" and in water are equivalent, if the measured amount is the same as the amount of air present in the lungs at the time weight-in-water is recorded" (p. 211).

Whether the measurement occurs on land or in water does not appear to have a substantial effect on the results, but it is apparent that measurements made on land are easier for the subjects, particularly if they are unaccustomed to water (144).

There are a number of techniques to measure RV, including nitrogen washout (178), helium (53,89) and oxygen dilution (8,36,93,134,150, 171); the method of choice usually depends on the facilities available. Body plethysmography is probably the most accurate method available, but few facilities possess the required apparatus (6). None of the gas dilution methods has a high degree of precision (144) and therefore some have suggested the use of a predicted residual volume, based upon measured vital capacity or anthropometric dimensions such as height or body mass (6,60,122,144). Bass (6), for example, recommended that RV was equivalent to 20% of the TLC in individuals between adolescence up to age 35 years, 25% from 35 to approximately 50 years, and 30-35% in individuals 50 to 70 years. Sinning (157) suggested that the use of vital capacity to establish residual volume provides little advantage over using a constant residual volume for all subjects. Lohman (110) estimated from Sinning's (157) data that an additional measurement error of approximately .003 g/cc is made when RVs are estimated, rather than measured.

Despite these sources of technical error, densitometry has been

demonstrated to be highly reliable, with a measurement error of .002 g/cc (144), or approximately 1% body fat (109,144).

In addition to technical measurement errors there are also biological errors associated with body density. The accuracy in determining body fatness is ultimately limited by biological variation in the composition of the FFB (38). Density of the FFB is known to change depending on age, racial background, gender, and degree of physical activity (110). This variation has already been briefly discussed, and will be more fully explained in the section on growth.

Lohman (110) has suggested that the biological variation in the FFB in a specific population is associated with an estimated variation in density of .006 g/cc. Combined with the technical measurement error of .002 g/cc (144), the total variation associated with body density is:

$$\sqrt{S_{biol}^2 + S_{tech}^2} = \sqrt{S^2_{total}}$$

$$\sqrt{(.006)^2 + (.002)^2} = \sqrt{.00004^2}$$

$$S_{total} = .0063 \text{ g/cc}$$

Lohman further suggests that this combined variance of .0063 g/cc, equivalent to 2.7% body fat, represents the inherent limitations of body density in accurately reflecting body fat content (110). However, Martin et al (118) found that, based on cadaver analysis, the SD of the density of the adipose tissue free mass was .02 g/cc, and therefore Lohman's (110) figures considerably underestimate the variability of the FFB and thus the limitations of body density in accurately reflecting body fat content.

A compilation of body density and FFB density values for males and females between the ages of 7-25 years can be found in Tables A1 to A4.

It is common practice to convert body density values into

estimations of percent body fat. As stated previously, in specific populations, this procedure is associated with a SEE equal to 2.7% body fat. The most commonly used equations are those of Siri (158),

$$\%F = (4.95/D - 4.5) \times 100,$$

which assumes constant densities of fat (.9 g/cc) and FFB (1.1 g/cc); and Keys and Brozek (100), later refined by Brozek et al (21),

$$\%F = (4.570/D - 4.142) \times 100,$$

which incorporates the concept of reference man and avoids the estimation of the density of the FFB. Within density units of 1.09 and 1.03, the two formulae agree within 1% fat (110). For subjects greater than 30% fat, the Siri formula predicts higher values than that of Brozek (110).

Lohman (108) demonstrated that the assumption of constant densities in children is invalid, and through his research and others (13,54,80,111) has estimated constants for individuals 7 to 25 years that allow for the derivation of age-specific equations to predict body fat from body density. For example, the derived equation for 9 to 11 year old boys is: $\%F = (5.3/D - 4.89) \times 100$.

The assumption of constant FFB density is also a problem with athletes, certain disease states (e.g., osteoporosis), and the elderly (17), and therefore predictions of percent body fat in these groups will be associated with a greater error than within those groups for which FFB density values are established. Thus, it is suggested to transform body density values into percent fat only when necessary (17, 138), and that these predictions should be given as ranges (e.g., 17 to 23%) rather than as specific values (e.g., 20%) (113).

In summary, although densitometry is considered to be the "gold

standard" method of body composition, it is only 10K gold. It is not without limitations and results should be viewed within the context of these limitations.

ii. Total Body Water

An alternative criterion method to densitometry is the measurement of total body water (TBW). TBW is the largest component of the body (24); approximately 60% of the average male weight is water, while slightly smaller proportions, approximately 50%, are noted in women (17). Neonates are composed of as much as 80 to 90% water (17,54,81, 166). By using an isotope dilution method, total body water can be measured relatively simply in populations that encounter difficulty with underwater weighing, such as the elderly and the young.

Isotopic dilution involves the administration of a known quantity of a water soluble isotope, and, after allowing time for equilibration (4,17,94,144,168), the measurement of the concentration of this isotope in a body fluid, such as serum, saliva, or urine. Since body fat is assumed to have a negligble water content (94), TBW can be used in conjunction with the water content of the FFB to estimate FFB mass, and by difference, fat mass (80,94,144,182).

Isotopic dilution techniques require 3 to 10 hours for isotopic equilibration (4,24,168,182). The two main substances used in isotopic dilution are deuterium oxide and tritiated water (17,24,80,144). Brodie (17) recommends that ideally the substance should be:

> "a) easy to use and preferably non-toxic; b) inert rather than radioisotopic; c) metabolised fairly slowly to allow time for data to be collected; d) evenly and rapidly distributed throughout the body and capable of achieving equilibrium as quickly as possible; e) accurate and convenient in the estimation and calculation of body composition values; f) cheap and readily available" (p. 33).

In choosing a substance, researchers should consider availability, ethics, cost, accuracy, validity, and limitations (17).

Although the measurement of total body water provides an alternative to densitometry, it is not without limitations. Firstly, the method is limited by the assumptions associated with the two compartment model. Secondly, circular reasoning is involved in the estimation of FFB mass, because a value for FFB density is required, which is partly determined by its water content (144). Thirdly, the percentage of water contained in the FFB must be known for an accurate estimation of body fat to be made. For example, if 73% of the FFB is considered to be water, then an estimation of FFB mass is derived by the following equation: FFB = TBW/.73 (4). It has been demonstrated that the proportion of water in the FFB decreases as aging occurs (54, 108,166) and thus unless an appropriate age-specific and population-specific constant is used, errors will result. Fourthly, individual isotopes are associated with specific limitations that have been outlined by Brodie (17). Finally, the accuracy of the technique is affected by the hydration level of the subjects under investigation (168). When body fat is predicted from isotope dilution, the associated measurement error is 1 to 2% body fat (109).

iii. Total Body Potassium

The third commonly accepted standard laboratory technique of body composition is the measurement of total body potassium (109), which provides an estimation of lean body mass (56). The estimation of total body potassium requires the use of a whole body scintillation counter to record the gamma ray emissions of the naturally occurring radioactive isotope of potassium, ^{40}K (4,17,24,43). The procedure

involves no hazard or discomfort to the subject other than claustrophobia (56). Based on the assumption that there is a constant ratio of potassium per kilogram of lean body mass (LBM), measurements of total body potassium can be used to calculate LBM and, by difference, fat mass (4,43). For example, 68.1 Meq is the value for the potassium content of the LBM as determined by early cadaver analysis (59). This constant can be used to estimate LBM via the following equation: LBM = TBK^+ / 68.1 Meq (56). However, one of the major limitations of this method is the fact that the K^+/LBM ratio is not constant (39,106,108); that it possibly varies between genders (17, 56), and does vary amongst specific populations (e.g., athletes, children, elderly) (17,24,108).

Further errors in the technique, cited by Smith and collegues (163), include an incorrectly assumed distribution of ^{40}K, differences in natural abundance, misapplied counting statistics, instability of counter performance, high background or instability of background, differences in body geometry, and variations in counting sensitivity. ^{40}K counts are poorly reproduced in fatter subjects (43, 56). Buskirk and Mendez (24) suggest that knowledge about total body potassium does provide additional information about body composition, but that there is uncertainty as to what components are being estimated, for cells of different types have different concentrations of potassium. The measurement error associated with the estimation of body fat by total body potassium is slightly larger, at approximately 3%, than that for TBW or densitometry (109).

The three standard laboratory techniques, densitometry, TBW, and total body potassium, are all based on the two compartment model, and

therefore are limited by the assumptions of this model. They are also time-consuming, require specialized equipment, and are not always suitable for use with the population of interest. The methods have been important, however, for validating clinically convenient techniques and new laboratory techniques.

iv. Other Laboratory Techniques

Other laboratory techniques have been designed to measure specific aspects of body composition, and these will be briefly discussed.

a. Muscle Mass Determinations

Skeletal muscle is the largest tissue in the body and consequently many investigators have attempted to measure it (109). Three common techniques are the measurement of total body nitrogen (4,17,24), creatinine excretion (3,11), and 3-methylhistidine excretion (17,24).

Total body nitrogen (TBN) is measured by neutron activation using gamma ray emission after irradiation (4,24). TBN provides a measure of total body protein, an indirect index of body cell mass (24). The method is often used in combination with the measurement of total body potassium to separate the muscle and non-muscle components of the LBM, through the technique of compartmental analysis. This procedure involves the use of constants for the K^+/LBM ratio and the N/LBM ratio of both muscular and nonmuscular tissue that may not always be generalizable to the population of interest (24). The ratios are affected by muscularity, body hydration, the protein and mineral content of the skeleton, and aging (17). The technique also involves a small dose of radiation and therefore is not suitable for all subjects (4).

The 24-hour urinary creatinine excretion is the most widely used

biochemical marker for the estimation of muscle mass (17). Creatinine is the only metabolite of creatine, 98% of which is located in the muscles and therefore an assay of 24-hour urinary creatinine excretion should theoretically reflect the total body creatine pool, and consequently muscle mass (144). The synthesis of endogenous creatine is influenced by the exogenous creatine in the diet; therefore a subject must be on a creatine-free diet, or a diet of a constant composition, for at least 3 days prior to the test (17,24). Other limits to the validity of the technique include strenuous exercise, emotional stress, the menstrual cycle, renal disease, and severe infection (109).

The method for 3-methylhistidine (3-MH) excretion is similar to that for creatinine, with the same errors associated with collection (17). The amino acid 3-MH is present almost exclusively in myofibrillar protein (17). During muscle protein breakdown the released 3-MH is neither recycled for protein synthesis nor metabolized further, but is excreted in the urine (17,24). Since muscle protein synthesis and degradation should be balanced processes during steady-state periods, daily urinary 3-MH excretion should be proportional to muscle mass in adult humans (24,121).

The 24-hour urinary collections must be complete and at least 3 days of a meat-free diet must be consumed before urinary collections are started (24). The main limitations of this method are that 3-MH does not represent all muscle protein breakdown (17,24), and the possibility of an incorrect assumption: that urinary 3-MH originates almost entirely from skeletal muscle (24). The technique is also limited by gender, age, maturation, nutrition, hormonal status, fitness, recent intense exercise, injury, and disease (24).

In general, the measurement of skeletal muscle mass by any of the techniques described is expensive, time-consuming, and confined primarily to research work (17,24).

b. Nuclear Magnetic Resonance Imaging

Nuclear magnetic resonance imaging (NMR), a technique used routinely in analytical chemistry, has shown promise as a body composition technique because it provides the opportunity to examine different tissues (18,82,109). NMR is based on the absorption or emission of electromagnetic waves by nuclei in response to particular magnetic fields (18,109). The technique requires a large bore magnet, capable of accepting a human body. A subject must lie as still as possible within the magnet for approximately 1 hour (18). The body is first exposed to a static magnetic field and then an alternating magnetic field of the same frequency. The measurement of the behaviour of hydrogen nuclei, mostly from water, after removing the attenuating magnetic field enables the formation of body images (18,109).

Although the static field, radiofrequency field, and alternating magnetic gradient fields are all possible hazards, there appears to be minimal harmful side effects associated with NMR imaging, and because it does not require radiation, it is considered to be a noninvasive technique (18).

NMR imaging offers tremendous promise in the study of various body compartments in health and disease (109), but at this time it remains cost prohibitive. Indeed, the full potential of NMR imaging as a body composition assessment tool may never be realized because of financial constraints (18). However, if this procedure can be used to validate other less expensive protocols, it could become the new "gold

standard".

c. Bioelectrical Impedance

Bioelectrical impedance is an electrical current injection method that measures total body resistivity (18,113,132,150). The method is based on the principle that impedance to the electrical flow of an injected current is related to the volume of a conductor (human body) and the square of the length of the conductor (height) (113,150). The measurement of impedance to a harmless 800 mA at 50 khz current is made after placement of electrodes on the wrist and foot of the subject. Current conduction is directly related to water content (18), and since lean tissue has a greater electrolyte and water content than fat, it possesses greater conductivity (7,113). Thus, bioelectrical impedance is an index of body water, and therefore fat free mass can be estimated from measurements of current resistance, and, by difference, fat mass (18,113).

The bioelectrical impedance methods are convenient, rapid, and noninvasive which makes them suitable for both clinical and research work (4,113,173). However, daily fluctuations in water content from exercise, dehydration, eating, and drinking all need to be standardized to obtain optimal impedance results (18,132). Khaled et al (101) reported a correlation of 0.98 between total body fat predicted from densitometry and that from impedance. Reported SEEs for prediction of percent body fat vary between 2.7 and 6.1%, which are similar to the more traditional body composition measurement techniques (18), although there appears to be systematic errors at the extremes of relative fatness (18,85,113,132). That is, there is a tendency to underestimate the relative fatness of obese individuals and overestimate the fatness

of lean individuals (85,113). Segal et al (150) suggest that the precision of predicting LBM from impedance can be enhanced by sex and fatness-specific equations. Malina (113) recommends further analysis of the effects of factors that may influence the electrolyte content of lean tissue, such as specific dietary components, bouts of acute physical activity, and certain disease states.

Deurenberg et al (41) have demonstrated that the bioelectrical impedance method is not able to detect relatively small changes in FFM that are associated with weight reduction, particulary if the changes are due to changes in the glycogen and associated water stores. After weight loss, the bioelectrical impedance method overestimates the FFM (41).

d. Total Body Electrical Conductivity (TOBEC)

TOBEC is also based on the principle that lean tissue conducts electricity better than fat does, but it is a current induction method (4,18,113), rather than a current injection method like bioelectrical impedance. TOBEC involves the use of a measurement chamber that contains a cylindrical coil large enough to accomodate a human subject (4,18,132). The coil induces an electromagnetic force which the subject disturbs as they pass through the coil. The size of the disturbance is proportional to the mass of conductive tissues and ion-containing fluids (18,132). The measure of conductance is then used to estimate FFB mass and by difference, fat mass (4).

Harrison (77) reported that pooled unpublished data on 158 subjects from a meeting of TOBEC users demonstrated a correlation of .984 between LBM estimated from a stature-adjusted TOBEC score and underwater weighing. TOBEC measurements are associated with

encouragingly low SEEs (approximately 3%) for predicting densitometrically determined % body fat (77). They have been correlated with TBW measurements or densitometrically determined body fat and FFB mass in infants (34, 51), adolescents (140), and adults (33), with correlations ranging from .877 to .997.

TOBEC takes 64 readings of FFB conductivity within a 10 second period (77,132). It is a safe, simple, convenient, rapid, noninvasive technique that does not require extensive operator training or recurrent expense (17,77,132). However, TOBEC readings are used to estimate TBW and subsequently FFB; therefore the same limitations that apply to TBW also apply to TOBEC. The main disadvantage is that of capital cost (17,18,77).

e. Ultrasonics/Radiographics

Ultrasonics and radiographic analyses have been used to assess muscle, bone and fat cross-sectional areas and to estimate total body composition (109). Ultrasound has been used to measure the depth of tissue interfaces, typically fat and muscle, and to produce images that allow for the examination of both superficial and deep structures (17). Ultrasound is a reliable, portable, noninvasive technique which may be particularly useful for measurement of obese individuals, where many problems are typically encountered (17,18).

A computed tomography (CT) scanner provides a cross sectional image of the distribution of x-ray attenuation, and thus an excellent cross sectional analysis of fat, muscle, and bone distribution can be obtained (18,57,109). CT scanning is particularly useful in the quantification of internal fat depots (151). Results from CT scanning have been compared to other body composition techniques, and it has

been demonstrated that anthropometric techniques typically overestimate the muscular area and underestimate the fat area in comparison (18, 57). Unfortunately, the capital cost and recurrent expense cost for servicing, film, magnetic tapes and discs are very high (18). Additionally, radiographic techniques, including x-ray and CT scanning, involve exposure to radiation, and therefore must be closely monitored (18,82). However, CT scanning allows for the detection of minute changes in body composition, and is particularly valuable in clinical situations where those changes may be significant, and in sports medicine research, to monitor atrophy and hypertrophy of muscle groups with injury, rehabilitation, and training (17).

f. Photon Absorptiometry

Photon absorptiometry is used to measure bone mineral content, usually in a small cross section of the forearm (109). There are two types: a single energy transmission, when 2 clearly defined tissues are involved; or a dual energy system, when variable thicknesses, such as in the trunk, are involved (17).

In the single photon method, a beam of photons (^{125}I) is passed over the forearm, and the amount which passes through is inversely proportional to the amount of bone mineral present (17,109). Early work suggested that total bone mineral could be estimated from a small cross section of bone; this method was therefore used in conjunction with other methods to study the body as a multicomponent system, rather than a two compartment system (109). However, it is unlikely that total bone mineral can be estimated from a small cross section of bone. In fact, the prediction to total skeletal mass from single photon absorptiometry has an error of 150-250 g (approximately 10%) in an individual, which

is large, and thus a measurement of total body calcium is often used (112).

Total body bone mineral content and LBM are measurable using dual photon absorptiometry, a technique which uses a whole body rectilinear scanner and a high activity (0.5-1.5 Ci) source of gadolinium-153 that emits energy at two discrete peaks (44 and 100 KeV) (112). Estimated precision of total bone mineral content measurements by dual photon absorptiometry for humans is high (2-3%) (112). This technique is portable, involves a low radiation dose (2-10 mrem), and allows for total body determination of bone mineral and LBM by relatively direct analysis (112). However, it's use may be limited by cost.

g. Infrared Interactance

Infrared interactance is a new method of body composition analysis based on light absorption and reflection principles (18,84). Using near-infrared spectroscopy, electromagnetic radiation is beamed against the site to be measured, and energy is either reflected, absorbed, or transmitted. The transmitted energy is scattered and reflected back out carrying with it information on the chemical composition of the material (84).

The method was initially used to determine the composition of grain and meat (18,84). It shows promise as a fast, safe, and noninvasive technique but it is still in the developmental stage (84).

h. Multicompartment Models

Investigators have recognized the limitations of the two compartment model, and have examined various multicompartment approaches to increase accuracy in the prediction of body composition. The primary limitation with the two compartment model is the assumption

of constant density and composition of the FFB. Water and mineral content are the most variable components of the FFB, and thus many multicomponent approaches include the measurement of both. Lohman (108,109) discusses the progression of multicomponent approaches, and provides equations for children for both a 3-compartment model, using body density, bone mineral, and percent fat; and a 4-compartment model, using density, water, bone mineral, and percent fat.

Lohman (108) suggests that theoretically, the 4-compartment model would enable estimation of percent body fat with an accuracy of less than 2% if technical errors in all measurements could be minimized. The 4-compartment model is ideal because the two most variable components of the FFB are directly measured. However, although body density can be measured in many laboratories, the measurement of body water and bone mineral content are not as routine (109). Practically, it appears that the 3-compartment model, including fat, density, and water, is the best multicomponent approach, particularly with the development of improved body water procedures (108). Although the 4-compartment model may be limited at this time, it may become more feasible in the future as techniques improve and costs drop (18).

2. Clinically Convenient Techniques

Because of the associated disadvantages of laboratory techniques, including cost, labor, inconvenience, and cooperation, more indirect but clinically convenient and noninvasive measures of adiposity have been developed. These include various skinfold indices, taken singly or together (4,27,31,43,108,138,146), and various height-weight indices, such as relative weight, body mass index (BMI), and ponderal index.

A. Skinfolds

The measure of skinfold thicknesses relies on laboratory techniques for validation, and therefore is considered to be a doubly indirect body composition method (30,31). Skinfold measurements are useful for assessing regional and overall adiposity and for monitoring fat loss for all age groups (26,30). They are used in physiology, structural anatomy, endocrinology, kinanthropometry, nutrition, health and fitness, growth, sport and exercise sciences (30). Skinfold thicknesses may be particularly useful during adolescence, where rapidly changing body composition makes height-weight indices unreliable (26,74). The accurate and reliable measurement of skinfolds, however, requires highly skilled personnel (26,74).

Skinfold measurements have been used in many ways to estimate adiposity. They have been used singly, as in the use of the triceps skinfold to indicate obesity (4,43,146); in combination, as in the Canadian Standardized Test of Fitness sum of 5 skinfolds (27); and to predict percentage of body fat, for which hundreds of equations have been derived (30,31,107,108,110,138,161).

Hager (73) emphasized two assumptions were necessary for the prediction of body fat from skinfold thicknesses: the first being that subcutaneous fat constitutes a constant proportion of total body fat over all ranges of body weight. The relationship between skinfold thicknesses and total body fat differs with ethnicity, age, and sex (12,29,79,110,161). Martin (117), as cited in Roche (144) demonstrated that the correlation between subcutaneous adipose tissue and total adipose tissue mass is approximately 0.8. The second assumption was that the sites of measurement are representative of all subcutaneous

fat. Mueller and Wholeb (128) and Mueller and Reid (130) have suggested that it is likely that a few skinfolds can represent all skinfold thicknesses, given the large proportion of variance in multiple skinfold thicknesses explained by a principal component in a multivariate analyses. However, other investigators (151,153,176) have demonstrated that there is a wide variance in fat distribution patterns, and therefore it is unlikely that one or two skinfold sites can represent all subcutaneous fat.

Clarys et al (30) elaborated further these assumptions following completion of the Brussels Cadaver Analysis Study, which was undertaken in order to validate body composition methods, particularly the skinfold measurement, skinfold formulae, the choice of skinfold site, and the relation among internal and subcutaneous adipose tissue depots. The study involved the analysis of 32 corpses, aged 55 to 94 years, with the exception of one aged 16 and one aged 33 years. Clarys et al (30) suggest 5 assumptions are required to derive body fat from skinfold thicknesses: 1. there is a constant compressibility in a double layer of skin plus subcutaneous adipose tissue; 2. skin thickness constitutes a negligble or a constant fraction of a skinfold; 3. there is a fixed adipose tissue patterning; 4. there is a constant fat fraction in adipose tissue; and 5. there is a fixed proportion of internal to external fat. In reality, there are problems with all of these assumptions.

In the Brussels study, the mean compressibility over all skinfold sites for each subject ranged from 38.2% to 69.3%, with a mean of 52.6%. There was no significant gender effect, but differences were observed at each site, with the front thigh and medial calf

demonstrating the lowest compressibilities and the supraspinale and biceps the highest. Skinfold compressibility, therefore, is not constant.

Skin thickness in each skinfold varied by site and gender. Skin thickness was greatest for the subscapular site in both males and females, while males generally had thicker skin than females. Clarys et al (30) suggest that while the contribution of skin to total skinfold thickness is generally not large, it may lead to significant error, especially in lean males. They further recommend that on the basis of skin thickness, the subscapular skinfold should be a poorer predictor than the skinfolds at arm and leg sites.

Many investigators have demonstrated variations in fat patterning (2,11,94,127,128,129,130,151,153,176). The two main concerns with regards to fat patterning and skinfolds are the number of sites that are required to be representative of the fat pattern, and the choice of the best sites. The best predictors of subcutaneous adiposity in the Brussels study were the front thigh, medial calf, rear thigh, and supraspinale. This supported the work of Mueller and Stallones (129), who suggested that a leg skinfold may be an important predictor of overall fatness, and was particularly important in describing the fat distribution pattern. The triceps skinfold, which many have reported to be the best single indicator of adipose tissue (4,12,17,43,146), ranked 11th out of the 14 skinfolds examined in the Brussels study.

The data from the Brussels study did not support the assumption of a constant fat fraction in adipose tissue. Limited data on 6 subjects demonstrated decreased water content and increased fat content with increased levels of adiposity. The level of fat in subcutaneous adipose

tissue also increases with age. Therefore, two identical thicknesses of adipose tissue may contain significantly different concentrations of fat.

The Brussels data demonstrated that skinfolds were significantly related to external (subcutaneous) adipose tissue, but the relation to internal tissue was less evident. The authors suggest that little can be said about the proportion of internal to external fat. However, others have suggested different age and sex trends in subcutaneous versus internal fat distribution. Generally, females have proportionately less subcutaneous fat and a higher proportion of internal adipose tissue mass per kg than males; while older subjects of the same sex have less subcutaneous fat than younger individuals in proportion to their total fat mass (17,110,151).

It should be noted that while the Brussels study provides new "direct" information on body composition, the information is limited to a sample of relatively older, sedentary individuals who demonstrated variation in body composition (e.g., adipose: 22.3 $\pm$ 7.3 kg; muscle: 21.3 $\pm$ 6.6 kg; bone: 8.5 $\pm$ 1.4 kg ($\bar{x} \pm SD$)). Some of the relationships that have been cited may in fact be different for a younger, active population. There is still a need for cadaver analysis to support and improve body composition methodology.

In addition to the biological problems that have been cited, there are a number of technical measurement problems associated with skinfold measurements. These include intra- and inter-tester reliability, tester experience, proper landmarking and definition of each site, the angle of the skinfold measurement, type of calipers used, and pressure exerted by the caliper (17,110,138). Skinfold measurements with the

obese are particularly difficult, as demonstrated in the following limitations outlined by Kuczmarski et al (104): 1. low accuracy in the amount of tissue picked up to form the skinfold; 2. difficulty in palpating the fat-muscle interface; 3. skinfold thickness may exceed the maximum opening of the caliper; 4. caliper tips may slide on larger skinfolds; 5. subsequent readings in obese subjects may be decreased due to repeated compression of subcutaneous fat; 6. differences in elastic properties of both fat and skin tissues between individuals and with age; 7. degree of compression involved in skinfold measures varies with location on the body; 8. variation in the depth at which the skinfold-caliper tips can be placed in obese persons; 9. interobserver variation; and 10. subject discomfort. Lohman's (110) review article contains a full discussion on the types of measurement errors that can occur with the measurement of skinfolds, and also with the estimation of body fat from skinfolds.

1. Prediction of Body Fat from Skinfolds

Many investigators have attempted to predict percent body fat from skinfold measurements. The common technique is to use regression analysis to derive a body density followed by the use of the Siri (158) or Brozek et al (21) equation. Therefore, the accuracy of these predictions depends not only on the accuracy of the estimation of density from skinfolds, but also on the validity of the assumptions underlying densitometry and the two compartment model (31). The estimation of body density or percent fat from skinfold measurements is convenient and quick, but it can be tremendously inaccurate and misleading when applied to a single individual, rather than a population (53).

a. Population-Specific Approaches

Early approaches in the prediction of body density or percent fat from anthropometric measures involved the development of population-specific regression equations. Research leading to the development of these equations demonstrated that age and gender are important sources of body density variation (138). Hundreds of specific equations were developed in the hope of improving the accuracy of predictions. This research approach suffered from a number of weaknesses: the use of small, homogeneous samples (138), which limit generalizability; the use of a large number of variables in comparison to number of subjects (38,53,87,93); the use of a linear model to describe the skinfold-density relationship, which is in fact curvilinear (12,86,108,110,138,171); the placement of too great an emphasis on stepwise multiple regression analysis, which minimizes the use of physiological theory and replaces it with a statistical edit (86,93,123,145,179); and a failure to cross-validate the prediction equations developed (38,86,93,110). The major limitation of the population-specific regression equations is their limited generalizability: they are affected by differences in age, gender, degree of fatness, and are derived from a linear, rather than a curvilinear model (86,138).

b. Generalized Approaches

Because of the limited generalizability of population-specific equations, researchers examined generalized equations for predicting body density from anthropometric measures (46,86,138). The generalized model extends the degree that an equation can be used within the accuracy defined by the standard error of prediction (86). Jackson (86)

suggests that the foundation for generalized body composition equations is derived from three research findings: 1. the skinfold/body density relationship is not linear in a heterogeneous sample, 2. age is independently related with body density, and 3. anthropometric variables are highly correlated. Generalized equations commonly address the flaws of population-specific equations. For example, a quadratic model is used to account for the nonlinearity between skinfolds and body density; age is used as an independent variable; and the sum of skinfolds is used rather than individual skinfolds, to maximize the subject/variable ratio (17,86,87,88). The generalized equations have not substantially improved the accuracy of the estimation of percent body fat; their advantage is that they are applicable to a larger group of people (17,38,86,138). They minimize the large prediction errors that occur at the extremes of the body density distribution that are associated with population-specific equations (138).

c. Estimation in Children

Many investigators have developed equations specifically for use in children (12,39,58,107,108,161,171,184). There are problems associated with the use of anthropometric variables to accurately and reliably predict body fatness in children (17,108). Lohman (108) has outlined four problems:

> "1. Equations derived for children and youth are often based on a wide range of prepubescent, pubescent, and postpubescent subjects. These equations have limited utility to each subsample;
> 2. Development of the regression equations is usually based on one body composition method and the two-component system, which provides biased estimates of fat content for chemically immature subjects if densitometry, hydrometry, or K^{40} spectrometry is used;
> 3. Few equations have been extensively cross-validated on other samples of children and youth and the identification of subsamples that may not yield accurate estimates also have not

been identified (eg., subjects of diverse athletic and racial background);
4. variation in the selection of sites and measurement procedures used for skinfolds, circumferences, and skeletal widths" (pp. 343-344).

Lohman (108) suggests that anthropometric equations for children need to be developed with careful consideration of all four problem areas.

d. Error

Skinfolds used in either population-specific or generalized equations predict body density with SEE equal to .0098 g/cc, which is close to that expected from biological and technical sources (110). Under optimal conditions, the SEE for percent fat predicted from anthropometric measures ranges from 3.5 to 3.9% (38,138). However, these figures do not include the biological error that is associated with the variability in density and composition of the FFB. According to Cureton (38), if it is assumed that the errors in determining body density and the error in predicting percent fat from body density are independent and additive, then the standard error for estimating the "true" percent fat can be estimated from the following equation:

$$\text{SE \%FAT} = \sqrt{\begin{array}{l}\text{Error Variance for} \\ \text{Estimating \%Fat} \\ \text{from Body Density}\end{array} + \begin{array}{l}\text{Error Variance for} \\ \text{measuring or} \\ \text{estimating body density}\end{array}}$$

If a value of 2.7% (110) is used as the SEE of converting body density to percent fat via the Siri equation, and it is assumed that the SEE for estimating body density is 3.5% (38,138), then the total error in estimating true percent fat is:

$$\text{SE \%FAT} = \sqrt{(2.7)^2 + (3.5)^2} = 4.4\%$$

The measurement or estimation of body density from anthropometric

measures is only part of the problem, for which a number of violated assumptions with regards to biological errors and technical sources of measurement error have been discussed. It appears that the accuracy in determining body fatness is ultimately limited by biological variation in the composition of the FFB (38,86).

e. Alternate Uses of Skinfolds

Because of the error associated with the prediction of percent fat from skinfold measurements, particularly within individuals, alternative uses of skinfolds have been suggested (17,27,30,82,93,113, 138,145,146). Many feel that skinfolds should be used directly, such as in a sum (27,58,93), or singly, as with the triceps skinfold (4,12), to give an idea of fat content, rather than introducing further error by attempting to derive total body composition. A sum of skinfolds provides a more representative sample of subcutaneous body fat and is more highly correlated with body density than individual sites (138). Sums of skinfolds or single skinfolds can be compared with national survey data (27,146) to derive a percentile rank for the individual and this rank can be interpreted as an indicator of percent body fat or total body fat and therefore an indirect grading of body fatness (146).

In summary, therefore, the measurement of skinfolds provides a quick, convenient technique for the assessment of body composition that can be reliable when done by trained personnel. Because there are errors associated with the estimation of percent body fat from skinfold measurements, it is wise to compare single subject skinfold values or sums of skinfolds to population norms to derive an indirect grading of body fatness.

B. Height-Weight Indices

Height-weight indices have been the choice of measurement for epidemiologists. The ease and accuracy with which height and weight can be measured provides a distinct advantage over other convenient measures such as skinfolds and circumferences. Height-weight indicators have been used to determine the degree of over- or underweight, in a variety of forms to indicate adiposity, and as an aid in health education (17,78,138).

i. Height-Weight Charts

There are three groups of height-weight indices. The first group consists of tables of average weights for height, age, and gender, such as those developed by Hamill et al (74) for the American National Center for Health Statistics. Relative weight is derived from tables of this type, and can be calculated as the weight expressed as a percentage of the standard weight for age, height, and gender (17,126,146), or the percentage of weight at the age when height is at the fiftieth percentile (149). Because these tables represent averages rather than ideal data, they are descriptive without providing standards (23,131).

The second group of indices are tables of "desirable" weights for height, such as the 1959 and 1983 Metropolitan Life Insurance Tables, which are based on the lowest mortality in an insured population at each weight for height level (23,78,131,155). These tables may be used for clinical, educational, informational, and research purposes (131). Tables of "desirable weights" are associated with a number of limitations. The "ideal" or lowest-risk weights will vary for different populations, and will vary for the same population at different times

and in relation to different causes of mortality (78). Therefore, an "ideal" or "desirable" weight cannot be identified for an individual who is different from the population on which desirable weight tables were based. Since the tables are based on insured populations, they are not representative and therefore applicable to the entire population, particularly lower socioeconomic and ethnic groups. The tables ignore other risk factors, such as smoking; they do not measure degree of fatness or fat distribution; they do not account for age; and they rely on an ill-defined concept of frame size (23,26,78,99,131,180). Often, the technical quality of the data used to establish the tables is variable, i.e., self-reported rather than measured (78). There are also sampling and design limitations in studies: in smaller studies the total mortality experience is too small to make a reliable interpretation of the association of risk with body weight, while larger studies are often not representative. Additionally, some studies use different units of analyses, such as insurance policies, rather than individuals, and therefore it is possible that the same person may be included a number of times in the analyses (78). The concepts of ideal or desirable weights may be helpful in dealing with population data, but they have limited utility when applied to the individual (25, 180).

It should be noted that there is a slight increase in "desirable" weights in the 1983 Metropolitan Life Insurance Tables, particularly in the shorter stature groups. It is now known that the suggested increased weights may contribute to disorders such as hypertension, hypercholesterolemia, and glucose intolerance, and therefore if using tables of "desirable" weights, the 1959 standards are recommended

(131).

Despite the noted limitations, height-weight charts also have a number of advantages. Firstly, height and weight are relatively easy to measure. Secondly, height-weight tables are ingrained in the North American culture. As Harrison (78) states "people in this culture want to know what they weigh, whether what they weigh is good or bad, and what they should do about it". Finally, evidence is increasingly demonstrating that high relative weight or obesity is related to an enhanced risk of premature death or disability (14,23,50,67,70,76,78, 105,115,131,155,156,172).

Since an ideal weight cannot be accurately identified for an individual at one point in time or for an individual who differs from the reference population, alternate strategies for height-weight indices have been proposed. These include: 1. shifting the focus to the ends of the weight continuum, and identifying undesirable or dangerous relative weights; 2. devising several appropriate tables for different groups in relation to various criteria; and 3. abandoning traditional height-weight tables in favor of a single index that does not require a reference population, such as body mass index (BMI) (78). It is from this third strategy that the third group of height-weight indices arise: this group includes those indices derived from height and weight, such as BMI (wt/ht^2), ponderal index ($wt^{.33}/ht$), Rohrer Index (wt/ht^3), Benn Index (wt/ht^p), and the Broca Index (wt/(cm - 100)).

ii. Body Mass Index (BMI)

A good index should be independent of height, and highly correlated with body weight and measures of body fat (99,147). The index which

best fits these criteria is BMI (17,62,69,99,138,145,147,155,156), which is the weight in kg divided by the height in metres squared. The BMI is correlated (r=0.7 to 0.8) with densitometrically determined body fat (14,17,144); is a highly significant predictor of total body fat (r=0.91) (18) and intra-abdominal fat (r=0.71) (18,124); is correlated with weight (r=0.89 to 0.98) (124); and is relatively independent of height (r=0.07) (17). BMI has been recommended by the Canadian Expert Group on Weight Standards (83) and the American National Institutes of Health Consensus Development Panel (26) as an appropriate measure of adiposity and associated health risk in the adult population. It has the advantages of ease and reliability of measurement, and is not based on a reference standard and therefore allows for comparisons between and within populations (26,115,131,156). However, the measurement lacks the ready interpretability of relative weight or desirable weight (115, 131). The primary limitations of the BMI are that it is stature dependent over at least part of the life cycle; it is affected by relative leg length or relative sitting height; and it reflects both lean and fat tissue to a comparable degree (65,123,146). Because only height and weight are used in the calculation, BMI predicts that every person of the same height and weight will possess the same amount of body fat (92), which is not true. One cannot lose sight of the individual differences in subjects with the same BMI. Ross et al (160) state that they "can show evidence where clinical evaluations based solely on the Quetelet or any weight for height index may be pure nonsense" (p. 18). BMI explains no more than 60 to 70% of the variability in body fat (62). It does not reflect regional distributions of body fat, or any age- or gender-related changes in

body composition. At any level of BMI, an older person will generally have proportionately more fat than a younger individual (76). Therefore, BMI should not be used with children or adolescents, pregnant women, or for predicting risk in the elderly (26,83), nor should it be used in isolation (27,83,123,135). The BMI is not good for use at the individual level, and as McLaren (119) states "Only when the BMI is abandoned in the study of obesity and more revealing and less misleading measurements are made will we begin to understand the different forms of obesity" (p. 121).

The use of BMI is enhanced with a measurement of body fat, such as a sum of skinfolds (27), and a measurement of health risk, such as the waist to hip ratio (83), as is done in the Canadian Standardized Test of Fitness (27).

Indices such as BMI can be improved for particular applications by generating population-specific exponents for height, as is done with the Benn Index (64). Different investigators, for example, have suggested that a power function of 1.5 is more appropriate for females than the normally used 2 (124,156). However, wt/ht^2 and wt/ht^p exhibit similar correlations with skinfold thicknesses, indicating that wt/ht^p is not a better indicator of adiposity (62). Additionally, wt/ht^p requires a series of cumbersome calculations such as age and sex-specific regression equations and transformations that are not easily done in clinical settings (62,64).

iii. Ponderal Index

The ponderal index ($wt^{.33}/ht$) is a generalized measure of linearity-laterality. It was originally developed on the basis of dimensionality theory: as linear dimensions increase, area varies as

the square of the linear dimension, and mass varies as the cube of the linear dimension (160). Unfortunately, the body form does not remain constant with increasing length. The ratio was inverted ($ht/wt^{.33}$) in the Sheldon somatotyping system (99,124), and this variation of the ponderal index has the disadvantage of being inversely proportional to weight at any given height (26,99,137). Higher ponderal index values indicate a more linear body build, while lower values indicate a more lateral build (152). The ponderal index is not as widely recommended as the BMI because it is not completely independent of height, and it is not as highly correlated with adiposity (17,99).

iv. Use of Height-Weight Indicators in Children

It is sometimes difficult to use weight for height measures in children. Obesity, for example, is often classified as a body weight more than 20% over the ideal weight, generally defined by using life insurance data for ideal weight (149); however, these data are not always readily available for children. Weight expressed as a percentage of weight at the fiftieth percentile for age correlates well with excess body fat, but fails to reflect increments in height or lean body mass that often accompany excessive adiposity in both children and adults (43).

The use of BMI in children and adolescents is limited as it varies with age (139,147) due to changes in body composition that occur with growth (4). Rolland-Cachera et al (147) developed normative curves for BMI for French children from a longitudinal growth study. They found that the correlation between height and BMI varied with age, being higher in children less than 1 year, and in those greater than 6 years. They related these changes to linear growth and adiposity growth. The

growth problem in the use of the BMI may be minimized by relating the measured BMI to the index of a theoretical child of the same age with weight and height on the fiftieth centiles of growth standards to give a percentage of BMI (139).

In summary, height-weight indicators are convenient and reliable. However, the three classes of indicators, including charts of relative weights, desirable weights, and indices formed of height and weight, are all associated with limitations. It is unlikely that any height-weight index will provide an acceptable indication of adiposity in the average range, but relationships become more apparent at the extremes of the distribution for each indice (99). Therefore, these indices may be used in the identification of obesity with some degree of accuracy. They should not, however, be used in isolation, but should be used in conjunction with other convenient indicators, such as skinfolds.

C. Visual Estimation

An alternative, simple method of body fat assessment is visual estimation (14,17,42,43,52,131,165,167). People rely on their vision to perform numerous tasks with great precision (167). Clinically, obesity is often diagnosed visually, but although such observations provide a good individual guide, they are not quantitative. Visual estimation should be combined with other convenient techniques, such as skinfolds or height-weight indices, to rule out subjectivity and possible cultural biases (42,52,126).

The visual diagnosis of obesity has rarely been used as a criteria in systematic studies because the measurement is hard to quantify. For example, Moynihan et al (126) found that visual ratings of obesity

differed from relative weight classifications of obesity in 10,932 British children. The authors suggested that the discrepancies found between the two methods demonstrate that one method of assessment cannot replace another. However, Sterner and Burke (167) demonstrated that some experienced raters can estimate body fat from photography almost as accurately as from skinfold measurements. The results of this study are promising, and if a visual technique for estimating percent body fat could be easily mastered, it would be less expensive and easier to administer than any of the commonly used field techniques. Thus, if proper validation studies are conducted, the visual estimation of body fat could be a well-recognized, convenient technique in the future.

3. Summary - Body Composition Methodology

In summary, the measurement of body composition is an inexact science in which one must be aware of the errors and shortcomings of the methods they are using. The three accepted standard laboratory methods, densitometry, total body water, and total body potassium, against which many laboratory techniques and all clinically convenient techniques are judged, are all based on the two compartment model. The primary limitation of the two compartment model is the assumed constant density and composition of the FFB, which has been demonstrated to vary with age, gender, disease, and activity. All three standard laboratory measures are associated with SEE of approximately 3-4% body fat. Newer techniques, such as NMR imaging, have shown promise but at this time are cost prohibitive. Other laboratory techniques, such as TOBEC, bioelectrical impedence, and infrared interactance need to be further validated before they are accepted as accurate and reliable measures of

body composition.

Convenient indicators, such as skinfolds and height-weight indices, have the advantage of quickness and ease of use, but are "doubly indirect" methods and therefore can only be as accurate as the criterion methods against which they are judged. Height-weight indices, such as BMI, which are highly correlated with measures of body fat as well as FFB mass, should not be used in isolation, but should be combined with a measure of fatness, such as skinfolds. The visual estimation of body composition may be a valuable tool in the future if it is possible to develop objectivity and remove cultural bias.

All techniques are associated with increased error when applied to the individual as opposed to a population. Therefore, if an estimation of percent fat is necessary, a range should be provided rather than a specific value to account for the error that is inherent within the specific technique. Otherwise, measures of body density or sums of skinfolds should be left in their "raw" form and compared to population standards to provide an estimate of an adiposity rating.

II. <u>What amount of fat is considered to be excessive?</u>

At the beginning of this review, obesity was defined as "an excess of body fat frequently resulting in a significant impairment of health" (23). There are a number of problems associated with the measurement of body fat. There is also a problem with dichotomizing a continuous variable (i.e., adiposity) into nonobese and obese categories. Obesity, like baldness, is a matter of degree. A certain amount of body fat is critical to survival, but at what point the accumulated energy reserve becomes clinically hazardous is unknown and is a subject of considerable debate (4,16,25,42,43,135,154,165,166). Indeed, not all

obese people are unhealthy (25). In adults, it has been demonstrated that marked obesity has adverse effects on health and longevity (14, 115,154,156,172), but controversy exists with regard to the moderately obese-mortality relationship in adults (115). Although antecedents of some adult diseases have been found in obese children, the mortality/morbidity relationships with adiposity, and therefore the health implications of obesity in children, are poorly established (42, 92,126,154). Consequently, it is unknown at what level increased fatness causes increased risk and therefore definitions of obesity are statistical, value-laden judgements rather than pathological in nature (43,135,154).

Similar criteria are often used for children and adults and for both gender, but there is little justification for this (126). The fact that there are changes in body fat and that these changes vary between gender suggests that criteria used to diagnose obesity should be age- and gender-specific (92). It is important to derive a correct, well-accepted definition of obesity, otherwise statements regarding prevalence, consequences, and the treatment of obesity do not make sense (68). For example, depending on the definition of obesity used, prevalence rates in children and adolescents vary between 3 to 40% (26,42,43,73,135,149). Recommendations for intervention and treatment of obesity in a population would obviously be very different if only 3% were identified as obese, versus 40%.

1. Statistical Definitions of Obesity

A number of convenient indicators have accepted classification points, while some indicators have more than one, as demonstrated in Table A.5. The individuals at the National Institutes of Health

Consensus Development Conference (131) recommended a criteria for obesity treatment of greater than 20% above desirable weight (23). This corresponds to a BMI of 27.2 and 26.9 for males and females, respectively, using the 1983 Metropolitan Life Insurance Tables, or BMIs of 26.4 and 25.8 using the 1959 Tables. All of these BMI figures fall into the Grade I obesity category of Garrow and Webster (69) and Bray (15), in which health risks of obesity begin to manifest themselves (90). These risks vary in degree of severity as a function of age, gender, and family history of dyslipidemia, hypertension, or diabetes (90).

The use of relative weight to define obesity is similar to desirable weight; both question the desirability of the normative data. For example, if a population is fatter than is desirable then any relative value may appear acceptable, even though the population itself is not (17). Also, a classification of obesity by relative weight must be interpreted with caution, as misclassification in some individuals is likely: those who have a high proportion of FFB mass and normal fat mass may be deemed obese, while those with a high percentage of fat but poor muscle development may be considered to be of "normal" weight (126).

The three grades of obesity classification developed for BMI represent moderate, massive, and morbid obesity (15,17,69,90,166). The health risks associated with obesity appear to begin with grade I obesity (90). Wilmore et al (180) use a slightly different definition for moderate obesity in females, suggesting a BMI range of 27 to 30, rather than 25 to 29.9. The Canadian Expert Group on Guidelines for Healthy Weights suggests there are two zones of risk: zone C, which is

equivalent to a BMI of 25 to 27, that may lead to health problems in some people; and zone D, which denotes a BMI of greater than 27 and is indicative of increasing risk of developing health problems (83).

The three National Center for Health Statistics (NCHS) surveys (NHES, NHANES I & II) defined overweight as ≥ 85th percentile of BMI for males (27.8) and females (27.3) between the ages of 20-29, who were studied between 1976-1980 (154). Severe overweight was defined as a BMI ≥ 95th percentile of the same group. Obesity and severe obesity were defined analogously, using the sum of triceps and subscapular skinfolds, rather than BMI. The rationale for the use of the 20-29 year old reference group as the standard is twofold: they were the leanest age group among adults studied; and virtually all age-related increases in body weight amongst adults are due to body fat accumulation (125, 154,172). It should also be noted that these are the criteria being used to monitor the American 1990 Objectives for the Nation regarding overweight (125).

As is noted in Table A.5, there are variations in classification points for obesity for percent body fat (14,17,93,107,150,166,177,180). Johnston et al (94) used ≥ 40% body fat as their criterion value for obesity in females, based on their observation that the borderline percentage of fat value for obesity in middle-aged females is approximately 37%. However, most authors agree that obesity should be defined as greater than or equal to 30% fat in females, and greater than or equal to 25% in males (14,17,93,166,180). Lohman (107) suggests that these figures represent mild obesity.

A triceps skinfold in excess of the 75th (1), 85th (42,43,135), and 90th (93,135) percentile for age and gender has been used to define

obesity. The 85th percentile value is one of the most widely accepted criterion for obesity, but there is no a priori justification for its use, except that statistically fatter members of the population are identified (43). Clarys et al (30) demonstrated that the triceps skinfold ranked 11th of 14 in terms of ability to predict subcutaneous adiposity. The correlation of the triceps skinfold with total body fat is not good when body fat level is high, primarily due to inaccuaracy and reproducibility of skinfolds in fatter subjects (43). Additionally, the narrowness of distributions of skinfold thicknesses in early childhood makes it difficult to diagnose obesity from skinfold measurements. For example, in 4 year old children the difference between the 5th and 95th percentile ranges from only 5 to 10 mm (92).

2. Comparative Studies

There are many authors who have compared different obesity criteria. Malina (113) suggests that in a cross-classification of the distributions of BMI and skinfold thicknesses, there will be some individuals who will be classified as overweight but not obese, i.e., a high BMI and low skinfold measure, and others who are classified as obese but not overweight. Garn et al (66) demonstrated with the Ten-State Nutrition Survey data that there were differences between the triceps skinfold and relative weight criteria: both indicators corresponded in adult women, while children deemed obese by the triceps skinfold criterion were less than 20% overweight, and obese adolescent girls considerably exceeded the conventional 20% overweight standard. They recommended that fatness should not be measured with relative weight indicators from infancy through to early adolescence. However, there are measurement problems with skinfolds in this age group, as

discussed previously.

Johnston (93) compared relative weight and skinfolds to a criterion percent body fat (males ≥ 25%; females ≥ 30%) in 169 males and 66 females, aged 12 to 17 years. His data suggested that in males, the triceps and subscapular skinfolds were better indicators of obesity than relative weight, while in females, there was no difference between skinfolds and relative weight. Johnston suggested the high relative weight values in males may reflect a larger lean body mass, and thus decrease the accuracy in obesity identification with relative weight. Johnston did not actually cross-tabulate various criteria measures, and therefore could not comment on whether the various indicators classified the same individuals as obese. However, in a later study using the same sample, Johnston (91) demonstrated that 31 males and 17 females were identified as obese by the %fat criterion. Of these, the triceps criterion correctly identified 15 males and 4 females, while the relative weight criterion identified 16 and 5, respectively. Johnston concluded that the indicators, at the currently used classification points, were low in sensitivity (23-50%) but high in specificity (85-100%).

Moynihan et al (126) demonstrated that different proportions of a large sample of British adolescents were classified as obese with relative weight and visual rating measures. Overall, 15.2% of the girls were visually estimated to be obese, compared with 8.9% identified by relative weight, while 7.8% were identified by both methods. In boys, 6.4% were visually identified as obese, 7.2% by relative weight, and 4.6% by both methods. The authors suggest that one method of assessment cannot replace another, and that the higher proportion of females

visually identified as obese may reflect different cultural attitudes towards obesity in males and females.

3. Measurement of Obesity in Children

There is considerable variance in the body composition of growing children at different ages in both gender, and therefore measures of obesity that are appropriate for adults are not applicable to children (26,92). Because the health risks of obesity have not been well identified in children, there is a further problem in identifying an appropriate classification point for the convenient indicators (42, 126). However, it has been well established that there is an increased risk for developing obesity in adulthood for persons overweight or obese at younger ages (29,92,107,135,185) and therefore the early identification of obese children is important. Lohman (107) has developed a fat and growth chart, and suggests that it is particularly important to monitor those individuals who increase percent body fat from one category to another as well as those who are in the high or very high categories.

The triceps skinfold and relative weight are the most widely used measures to assess obesity in children (43). Both are associated with limitations. Zack et al (185) stress that studies of adolescent obesity should use a measure of adiposity such as skinfolds, rather than height-weight indicators, which have been demonstrated to be unreliable in children. This has been supported by others (66,93). In practice, most obese children are identified by their appearance.

In summary, the definition of obesity is currently a value-laden, judgement decision, rather than one which has a pathological basis. A number of convenient indicators have accepted obesity classification

points, but these points are not based on common percentiles, and therefore classify individuals differently. The acceptance of a quantifiable definition of obesity is important for the accurate establishment of prevalence statistics and the organization of treatment programs.

4. Health Risks of Obesity

Although not all obese people are unhealthy, obesity has been related to increased mortality and morbidity. Obesity has been associated with an increased prevalence of diabetes (14,37,92,137,155, 167,172,180), hypercholesterolemia (37,167,172,180), hypertension (14, 37,92,137,167,172,180), gall bladder disease (14,15,37,92,155); endometrial, uterine, cervical, and breast cancer in women (14,15,92, 131,137,154,167); prostate, rectal, and colon cancer in men (14,15, 131,137,154,167); coronary artery disease (92,131,137,154,155,167,180); abnormalities in pulmonary function (14,15,37,70); mechanical problems (15); and psychological problems (37,92,131).

Health risks are independently related to the amount of fat an individual possesses, as well as to the location of the fat (2,4,5,11, 14,18,90,95,115,176,180). A male or android fat pattern of abdominal, centralized, or upper body obesity, commonly known as the apple or "spare tire", is independently associated with hyperinsulinemia, diabetes mellitus, hypertension, and hypertriglyceridemia (5,11,14,90, 95,180). Fat patterning is commonly measured with the waist to hip ratio (2,4,5,11,14,26,95,176), with values of greater than or equal to 1.0 in males and 0.8 in females suggested as being indicative of risk (26).

A number of epidemiological studies have attempted to define the

relationship between body weight or BMI and mortality. This relationship has been described as linear (14,52,115), J-shaped (15, 115,155,156), and U-shaped (14,15,51,70,115,137,155,156), with optimal weights or BMIs varying between populations, age groups, gender, and time frames. The most recent information suggests that the obese, or subgroups of the obese, have increased health risks, but in the seventies a slight degree of overweight was considered healthy. Andres (3) concluded that

> "it would appear that for healthy individuals the weights associated with minimal mortality increase with age and that the old saw that one's best weight is that achieved at age 20 or 25 needs to be discarded " (p. 318).

Keys (98) concluded

> ". . the idea has been greatly oversold that the risk of dying prematurely or of having a heart attack is directly related to the relative body weight. For middle-aged men, the best prospect for avoiding death in ten or 15 years is to be about average, or a bit over, in relative weight" (p. 306).

However, a 26-year follow-up of the Framingham Heart Study population indicated that weight gain after the young adult years conveyed an increased risk of cardiovascular disease in both sexes (67,155). Clearly, controversy exists in the interpretation of the weight-mortality relationship.

Manson et al (115) examined 25 major prospective studies on the relationship between body weight and longevity, and concluded that each study suffered from at least one of three major biases: 1. failure to control for cigarette smoking; 2. inappropriate control of biologic effects of obesity, such as hypertension and hyperglycemia; and 3. failure to control for weight loss due to subclinical illness. Cigarette smoking is an independent risk factor for mortality and is also strongly associated with lower body weights, and therefore failure

to control for its effects will produce an artifactually high mortality in lean subjects (115,116). It is incorrect to statistically control for conditions such as hypertension and diabetes since obesity may adversely affect health by increasing morbidity as well as the case fatality rate; total mortality is an integrated function of both processes (115,167). Clinical and subclinical illness can reduce body weight and artifactually increase the mortality due to lower weights, particularly in studies with short follow-up periods (115,154,167). The length of follow-up must be carefully considered in interpreting the weight-mortality relationship (154,156). The effect of obesity on mortality is delayed, so that it is not seen in short term studies (2, 154,155,156). Over sufficiently long follow-up periods, e.g., 15 to 20 years, even a moderate degree of overweight confers excess risk for disease (2,105). Additional features in epidemiological studies which may result in discrepant findings include small sample sizes, inconsistencies in the obesity index used (115,164), and variations in research methodology and analytical techniques (164). Manson et al (115) suggest that the three largest studies they examined, the 1959 Build and Blood Pressure Study, the 1979 Build Study, and the American Cancer Study, reveal similar weight-mortality curves: the shape is closer to a J than a U, and a completely linear relationship is plausible since biases from smoking and selective early mortality have not been accounted for. Mortality at average weights is at least 5 to 10% higher than mortality at "desirable" weights. Simopoulos and Van Itallie (156) have also suggested, after reviewing the data from the 3 previously mentioned studies and the Framingham and Manitoba studies, that the weights associated with the greatest longevity tend to be

below the average weight of the population under consideration, if such weights are not associated with a history of significant medical impairment. Lew (105) and Stallones (164) also concur with these suggestions.

5. Health Risks of Obesity in Children

Childhood and adolescent obesity is related to decreased growth hormone release (20,75,141), hyperinsulinemia (2,120,143), carbohydrate intolerance (20,43), hypertension (153), hyperlipidemia (8,49,120), altered adreno-cortical function (75), and orthopedic problems (43). Functionally, obese youngsters have low fitness levels, as indicated by physical work capacity measures and maximal oxygen consumption measures (48), a decreased exercise and heat tolerance (141), and a decreased respiratory ability (120,165).

The psychological and social hazards associated with the early onset of obesity may be more debilitating than the medical hazards (20, 42,92,135,165). Obese children and adolescents demonstrate greater body image disturbances (42,143), poorer self-concepts, and more evidence of psychological dysfunction than their normal weight peers (42). Obesity has a low social value in general, which is reflected in less acceptance from peers (19,141), discrimination against the obese in school, admission to college, and other related activities (135), including the second rate medical care they are liable to receive (139).

III. Growth and Obesity

The composition of a child's body changes as a result of the processes of growth and development and consequently these processes should be considered in the study of childhood and adolescent obesity.

This brief review will concentrate on the body composition changes that occur with growth.

Obesity has been viewed as a disturbance of a normal pattern of growth. Forbes (56) identified two types of childhood obesity in a relatively small sample of 27 subjects: 1. developmental, which was characterized by an increase in lean body mass, fat mass, and stature, a more rapid rate of maturation, and likely presence since infancy; and; 2. reactive, which was associated with an increase in fat mass, and a more recent onset. Others have confirmed that obese children frequently are tall for their age, possess excess lean tissue as well as fat tissue, and mature early (43,135,148,165,175), but not all obese children display these characteristics (92).

1. Growth of Fat Tissue

The growth of fat tissue has been commonly assessed via the use of skinfolds in various population surveys. The limitations of the use of skinfolds have been discussed previously, and therefore the results reported, particularly those in young children, must be viewed with caution. Harsha et al (79) utilized data on a cohort of 447 children from the Bogalusa Heart Study that had been monitored yearly from birth or 6 months until 7 years of age. They determined that the subscapular skinfold decreased from 6 months to ages 4 and 5 in all race-sex groups, and thereafter increased again in all groups to age 7, except in black boys. Garn and Clark (66), using triceps skinfold data from the Ten-State Nutrition Survey, demonstrated that females possessed greater subcutaneous fatness at all ages from childhood through the ninth decade; that there was a prepubertal fat gain in both gender; but an adolescent fat loss in the male coincident with an adolescent fat

gain in the female. Johnston (91), relaying triceps skinfold data on 2 to 18 year olds from the Health and Nutrition Examination Survey I, confirmed the findings of Garn and Clark. He observed that at every age and percentile, females have thicker subcutaneous fat than males, that males show a decrease in skinfold thickness between 13 to 17 years, and that the range of variation in skinfold thickness, while limited in childhood, increases steadily thereafter. These findings have been confirmed by others (43,58,93,127,135,184).

Differences between gender become obvious during the pubertal growth spurt: there are changes in the amount as well as in the distribution of fat. Chumlea et al (29) found that there was little or no change in percent body fat in girls between 10 and 18 years of age, while males demonstrated a mean annual decrease of 1.15%/year for the same time period. However, males demonstrated no significant change in the amount of total body fat, while females showed a mean annual increase of 1.14 kg/year. Mueller (127) found that during adolescence and the third decade of life a redistribution of body fat occurs, with fat moving away from the extremities towards the trunk. The process occurs in both gender, although more rapidly in males than females.

Riumallo and Durnin (142) studied body composition changes through the use of 4 skinfolds in 47 boys at age 10 and 13 years. They proposed a two-stage growth process that can be characterized firstly by a "fat wave", where boys gain weight that is approximately 40% fat, followed by a "growth spurt", where weight gained is almost 100% lean mass, while using up some of their body fat. The timing, magnitude, and direction of the fat and growth spurt is variable, and may be modified by environmental factors, particularly nutrition and physical activity

levels. Other investigators have demonstrated the presence of the fat wave in females (58,66), but girls tend to lay down progressively more fat as time goes on, while in males the fat wave is temporary. The maximum fat deposition for both genders occurs approximately 2 years before peak height velocity (135). Knowledge of this fat wave is important in the body composition measurement of young boys and girls, as it may be possible that obesity classification occurs as a result of it.

2. Growth of Lean Body Mass

Lean body mass, estimated by ^{40}K counting in 609 subjects aged 7.5 to 20.5 years of age, demonstrated changes with age according to curves which closely parallel height and weight (55). Prior to puberty, there is no appreciable gender difference in lean body mass. Generally, during the pubertal growth spurt males have an increase in LBM and a decrease in fat mass, while females have a more gradual increase in LBM and an increase in fat mass. The maximal increase in muscle mass occurs at the time of peak height velocity in both gender (135,136).

To accurately investigate body composition in growing children, particularly FFB composition, knowledge of the potassium, protein, mineral, and water content is necessary. Excellent reviews have recently been provided on the FFB composition in children (13,29,54, 81,108,111,161).

1. Water Content of FFB

Children have a higher total body water and higher water content in the FFB (about 76%) than adults (72-73%) (13,108). Boileau et al (13) examined the water content of the FFB in 292 males and females aged 8 to 30 years, and found that the percent of water/FFB progressively

decreased from prepubescence to adulthood by 2.8% overall, at a rate of 0.38% per year for both genders, with the lowest values found in the adult sample. The water content of the FFB was lower for males (73.5%) than females (74.2%). Lohman (108) suggests that the use of adult water constants for estimating fat from density and other indirect body composition methods leads to an overestimation of the FFB and an underestimation of fat content.

ii. Mineral Content of FFB

Haschke et al (81) and Fomon et al (54) estimated the mineral content of the FFB in reference children from birth to age 10 years, based on the photon absorptiometric measurement of the forearm to derive total body calcium using data from the literature, assuming that the calcium content of osseous mineral is 34% in children and adults. They demonstrated a mineral content of 3.7% in male and female fetuses and infants which gradually increases to the adult value of 6.8%. Lohman and coworkers (111) also used photon absorptiometry to estimate the bone mineral content and bone width of the radius and ulna in 292 children and adults. They estimated the mineral contents for the prepubescent child, adult male, and adult female to be 5.2%, 6.6%, and 6.0%, respectively. Generally, therefore, it appears there is a low mineral content in the FFB in childhood that progressively increases until adult values are reached. It appears that the 18-year old male is within 95 to 98% of the adult mineral content (108).

iii. Potassium Content of FFB

A gradual increase in potassium content is evident in boys from birth to 18 years of age (54,81,108). Lohman (108) demonstrated that the K^+/FFB ratio in children and youth is 2.5 g K^+/kg FFB, which is

lower than the reported adult male value of 2.66 g K^+/kg FFB, but similar to the adult female values of 2.3 to 2.5 g K^+/kg FFB. The use of adult potassium constants to estimate percent fat in children and adolescents will result in an overestimation of body fat (81).

iv. Protein Content of FFB

The protein content of the FFB gradually increases from 19% in the prepubescent male child to 20.3% in the 18-year old adolescent (54,81, 108). Protein estimates are based on the body potassium data and on the assumption that a constant nitrogen:potassium ratio of 0.469 g/mmole is characteristic of children and youth (81). This assumption requires further validation over a broad age range (108).

In summary, it is apparent that as a child grows, a chemical maturation in the composition of the FFB occurs: there is a decrease in the percentage of water, and an increase in the percentage of bone mineral, potassium, and protein content with maturation. Chemical "immaturities" in younger children and adolescents result in a FFB density which is less than 1.1 g/cc (39,80,110,111). Therefore, the use of adult constants in denistometry will result in an overestimation of percent fat in children (107). Lohman (108) suggests that chemical maturity occurs between 15 to 18 years of age. Prior to this age range, it is imperative that age- and gender-specific constants be used to improve the accuracy of body composition estimations.

IV. Principal Components Analysis

Principal components analysis (PCA) is a statistical technique that may be applied to a group of variables in which none has been specified as the dependent of independent variable. PCA differs from other multivariate techniques in that intercorrelations among variables in a

single set, as opposed to external criteria, are analyzed (169).

In PCA, the investigator is usually interested in discovering which variables in a data set form coherent subgroups (i.e., factors or components) that are relatively independent of one another. As Tabachnick and Fidell (169) state:

> "The specific goal of analysis may be to summarize patterns of intercorrelations among variables, to reduce a large number of variables to a smaller number of clusters while retaining maximum spread among experimental units, to provide an operational definition (a regression equation) for an unobserved, hypothetical construct by using observed variables, or to test a theory about the nature of underlying variables." (p. 372)

There are two major uses of PCA: exploration and confirmation. Exploratory PCA, usually performed early in the research process, seeks to reduce data by grouping together variables that are intercorrelated (133,169). The variables themselves may or may not have been chosen with potential underlying structure in mind (133,169). Confirmatory PCA, usually performed later in the research process, is used to test hypotheses about the underlying structure of the data.

Steps in PCA (169) include:

> ". . . selecting and measuring a group of variables, preparing the correlation matrix, determining the number of components or factors to be considered, extracting a set of components or factors from the correlation matrix, (probably) rotating the components or factors to increase interpretability, and, finally, interpreting the results. Although there are statistical considerations relevant to most of these steps, the final test of the value of an analysis usually depends on the interpretability of the results: a good PCA or FA "makes sense"; a bad one does not." (p. 373)

With PCA, a set of n correlated real variables is replaced by a set of uncorrelated variables (principal components), which are linear transformations of the original variables:

$$L = b_1 l_1 + b_2 l_2 + b_3 l_3 + \ldots + b_n l_n$$

where b_1, b_2, etc. are weights chosen to maximize the explained variance (128,130). As Mueller and Reid (130) explain:

> ". . The extent to which a single principal component, L, can explain the bulk of the variance depends on the extent to which the n original variables are correlated. If correlations are high and of similar direction, the first principal component may explain almost all of the variance". (p. 20).

There are a total of n possible principal components from an analysis of n variables. Subsequent components will be independent of the first and one with another, and each explains less of the multivariate variance than the one preceding it (128,130).

For each component, a score (i.e., factor score) can be computed for each individual based upon the weights obtained from the PCA. These scores represent scores subjects would have received had the component been measured directly (169). Factor scores are estimates only, and changes in extraction or rotation may change the estimate considerably (169).

TABLE A1: BODY DENSITY VALUES IN BOYS

AGE	N	DENSITY ($\bar{x} \pm SD$)	METHOD	REFERENCE
8	9	1.055±.011	underwater weighing	39
9	17	1.051±.015		
10	12	1.054±.017		
11	11	1.052±.009		
10	40	1.057±.014	underwater weighing	136
11	40	1.059±.018		
12	40	1.062±.015		
13	40	1.070±.014		
14	40	1.066±1.012		
15	40	1.065±.013		
16	40	1.071±.011		
17	40	1.076±1.012		
10.5		1.051	obtained from calculated densities of FFB mass	80
11.5		1.048		
12.5		1.051		
13.5		1.055		
14.5		1.058		
15.5		1.061		
16.5		1.063		
17.5		1.064		
18.5		1.064		
17.43	141	1.0798±.0096	underwater weighing	171
9.9	42	1.049±.017	underwater weighing	161
12.5	23	1.056±.019		
15.5	60	1.064±.015		
10.1	31	1.045±.014	underwater weighing	13
12.3	11	1.050±.022		
15.5	34	1.061±.015		

TABLE A2: FFB DENSITY VALUES IN BOYS

AGE	FFB DENSITY	REFERENCE
1 mo	1.064	54
1 yr	1.068	
2	1.072	
3	1.074	
4	1.076	
5	1.078	
6	1.080	
7	1.081	
8	1.082	
9	1.084	
10	1.085	
10.5	1.086	80
11.5	1.086	
12.5	1.086	
13.5	1.087	
14.5	1.088	
15.5	1.090	
16.5	1.091	
17.5	1.092	
18.5	1.093	

TABLE A3: BODY DENSITY VALUES IN GIRLS

AGE	N	DENSITY ($\bar{x} \pm SD$)	METHOD	REFERENCE
9-10	19	1.0413±.0153	underwater weighing	184
11	13	1.0337±.0122		
12	21	1.0336±.0142		
13	14	1.0380±.0154		
14	14	1.0366±.0076		
15	9	1.0443±.0136		
16	12	1.0345±.0090		
9.9	20	1.036±.017	underwater weighing	161
11.6	29	1.035±.017		
15.2	58	1.041±.013		
10.1	12	1.038±.018	underwater weighing	13
12.1	17	1.035±.018		
15.3	29	1.041±.012		
16.5	133	1.0661±.0105	underwater weighing	171

TABLE A4: FFB DENSITY VALUES IN GIRLS

AGE	FFB DENSITY	REFERENCE
1 mo	1.064	54
1 yr	1.069	
2	1.071	
3	1.071	
4	1.072	
5	1.073	
6	1.073	
7	1.073	
8	1.074	
9	1.074	
10	1.075	

Table A.5: Acceptable Classification Points for Obesity using Convenient Indicators

Indicator	Classification Point Male	Classification Point Female	Reference
% Body Fat	≥20%	≥30%	150
	≥25%	≥30%	14, 17, 93 166, 180
		≥32%	177
		≥40%	107
Triceps Skinfold (for age, gender)	≥75th percentile		1
	≥85th percentile		42, 43, 66, 135
	≥90th percentile		93, 135
Weight for height, age, gender	≥85th percentile		23
	≥90th percentile		135
	≥95th percentile		146
	≥97th percentile		149
Relative Weight	≥ 120%		14, 16, 17, 23, 42, 93, 126, 135, 149, 165
Desirable Weight	≥ 120%		23, 37, 131
Relative BMI	≥ 120%		139
BMI	grade I -25 - 29.9		
	grade II -30 - 40		69, 90, 166
	grade III - > 40		

References

1. Adeyanju M, Creswell WH, Stone DB, Macrina DM. A three-year study of obesity and its relationship to high blood pressure in adolescents. J School Health 1987;57(3):109-113.

2. Anderson AJ, Sobocinski KA, Freedman DS, Barboriak JJ, Rimm AA, Gruchow HW. Body fat distribution, plasma lipids, and lipoproteins. Arteriosclerosis 1988;8:88-94.

3. Andres R. Mortality and obesity: the rationale for age-specific height-weight tables. In: Andres R, Bierman EL, Hazzard WR, eds. Principles of geriatric medicine. New York: McGraw-Hill Book Company, 1985:311-318.

4. Bandini LG, Dietz WH. Assessment of body fatness in childhood obesity: evaluation of laboratory and anthropometric techniques. Am Dietit Assoc 1987;87(10):1344-1348.

5. Barrett-Connor EL. Obesity, atherosclerosis, and coronary artery disease. Ann Intern Med 1985;103(6 pt 2):1010-1019.

6. Bass BH. Functional residual capacity and residual volume. Lung Function Tests. London: HK Lewis & Co. Ltd. 1974:27-31.

7. Baumgartner RN, Chumlea WM, Roche AF. Associations between bioelectrical impedance and anthropometric variables. Hum Biol 1987;59(2):235-244.

8. Becque MD, Katch VL, Rocchini AP, Marks CR, Moorehead C. Coronary risk incidence of obese adolescents: reduction by exercise plus diet intervention. Pediatrics 1988;81(5):605-612.

9. Behnke AR, Wilmore JH. Evaluation and regulation of body build and composition. Englewood Cliffs, NJ: Prentice-Hall, 1974.

10. Berkowitz RI, Agras WS, Korner AF, Kraemer HC, Zeanah CH. Physical activity and adiposity: a longitudinal study from birth to childhood. J Pediatrics 1985;106(5):734-738.

11. Bjorntorp P. Regional patterns of fat distribution. Ann Intern Med 1985;103(6 pt 2):994-995.

12. Boileau RA, Wilmore JH, Lohman TG, Slaughter MH, Riner WF. Estimation of body density from skinfold thicknesses, body circumferences and skeletal widths in boys aged 8 to 11 years: comparison of two samples. Hum Biol 1981;53(4):575-592.

13. Boileau RA, Lohman TG, Slaughter MH, Ball TE, Going SB, Hendrix MK. Hydration of the fat-free body in children during maturation. Hum Biol 1984;56(4):651-666.

14. Bray GA. Overweight is risking fate. Definition, classification, prevalence, and risks. Ann NY Aca Sci 1987;499:14-28.

15. Bray GA. Complications of obesity. Ann Intern Med 1985;103(6 pt 2):1052-1062.

16. Brennan N, Kevany J. Anthropometry and body image in a selected sample of adolescent girls. IJMS 1985;154(6):220-227.

17. Brodie DA. Techniques of measurement of body composition. Part I. Sports Med 1988;5:11-40.

18. Brodie DA. Techniques of measurement of body composition. Part II. Sports Med 1988;5:74-98.

19. Brownell K. The psychology and physiology of obesity: implications for screening and treatment. Am Dietit Assoc 1984;84:406-414.

20. Brownell K, Kaye, F. A school-based behaviour modification, nutrition education, and physical activity program for obese children. Am J Clin Nutr 1982;35:277-283.

21. Brozek J, Grande F, Anderson JT, Keys A. Densitometric analysis of body composition: revision of some quantitative assumptions. Ann NY Aca Sci 1963;110:113-140.

22. Brozek J, ed. Body composition. Ann NY Aca Sci 110, Volumes I and II, 1963.

23. Burton BT, Foster WR, Hirsch J, Van Itallie TB. Health implications of obesity: an NIH Consensus Development Conference. Int J Obesity 1985;9:155-169.

24. Buskirk ER, J Mendez. Sports science and body composition analysis: emphasis on cell and muscle mass. Med Sci Sports Exer 1984;16(6):584-593.

25. Callaway W. Weight standards: their clinical significance. Ann Intern Med 1984;100(2):296-298.

26. Canadian Dietetic Association. Obesity: a case for prevention. Official position of the Canadian Dietetic Association. J Can Dietit Assoc 1988;49(1):11-16.

27. Canadian Standardized Test of Fitness (CSTF) Operations Manual. 3rd ed. Ottawa: Minister of State, Fitness and Amateur Sport, 1986.

28. Cheek D. Body Composition Studies (letter to the editor). Am J Clin Nutr 1980;33:2060-2061.

29. Chumlea WC, Siervogel RM, Roche AF, Webb P, Rogers E. Increments across age in body composition for children 10 to 18 years of age. Hum Biol 1983;55(4):845-852.

30. Clarys JP, Martin AD, Drinkwater DT, Marfell-Jones MJ. The skinfold: myth and reality. J Sports Sciences 1987;5:3-33.

31. Clarys JP, Martin AD, Drinkwater DT. Gross tissue weights in the human body by cadaver dissection. Hum Biol 1984;56(3):459-473.

32. Cochran WJ, Fiorotto ML, Sheng HP, Klish WJ. Reliability of fat-free mass estimates derived from total-body electrical conductivity measurements as influenced by changes in extracellular fluid volume. Am J Clin Nutr 1989;49:29-32.

33. Cochran WJ, Wong WW, Fiorotto ML, Sheng HP, Klein PD, Klish WJ. Total body water estimated by measuring total-body electrical conductivity. Am J Clin Nutr 1988;48:946-950.

34. Cochran WJ, Klish WJ, Wong WW, Klein PD. Total body electrical conductivity used to determine body composition in infants. Pediatr Res 1986;20:561-564.

35. Conway JM, Norris KH, Bodwell CE. A new approach for the estimation of body composition: infra-red interactance. Am J Clin Nutr 1984;40:1123-1130.

36. Cote RW, Wilmore JH. A practical assessment of body composition in young women. J Sports Med 1986;26:427-430.

37. Council on Scientific Affairs. Treatment of obesity in adults. JAMA 1988;260:2547-2551.

38. Cureton KJ. A reaction to the manuscript of Jackson. Med Sci Sports Exer 1984;16(6):621-622.

39. Cureton KJ, Boileau RA, Lohman TG. A comparison of densitometric, Potassium-40 and skinfold estimates of body composition in prepubescent boys. Hum Biol 1975;47(3):321-336.

40. Deurenberg P, Westrate JA, Hautvast JG. Changes in fat-free mass during weight loss measured by bioelectrical impedance and by densitometry. Am J Clin Nutr 1989;49:33-36.

41. Deurenberg P, Westrate JA, Van der Kooy K. Body composition changes assessed by bioelectrical impedance measurements. Am J Clin Nutr 1989;49:401-403.

42. Dietz WH. Childhood obesity: susceptibility, cause, and management. J Pediatrics 1983;103:676-685.

43. Dietz WH. Obesity in infants, children, and adolescents in the United States. I. Identification, natural history, and aftereffects. Nutr Research 1981;1:117-137.

44. Donnelly JE, Brown TE, Israel RG, Smith-Sintek S, O'Brien KF, Caslavka B. Hydrostatic weighing without head submersion: description of a method. Med Sci Sports Exer 1988;20(1):66-69.

45. Durnin JV, Lonergan ME, Good J, Ewan A. A cross-sectional nutritional and anthropometric study, with an interval of 7 years, on 611 young adolescent schoolchildren. Br J Nutr 1974;32:169-179.

46. Durnin JV, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged 16 to 72 years. Br J Nutr 1974;32:77-92.

47. Durnin JV, Rahamin MM. The assessment of the amount of fat in the human body from measurements of skinfold thickness. Br J Nutr. 1967;21:681-689.

48. Epstein L, Koeske R, Zidansek J, Wing R. Effects of weight loss on fitness in obese children. Am J Dis Child 1983;137:654-657.

49. Epstein L, Wing R, Valoski A. Childhood obesity. Pediatr Clin of North Am 1985;32(2):363-379.

50. Feinleib M. Epidemiology of obesity in relation to health hazards. Ann Intern Med 1985;103(6 pt 2):1019-1024.

51. Fiorotto ML, Cochran WJ, Klish WJ. Fat-free mass and total body water of infants estimated from total body electrical conductivity measurements. Pediatr Res 1987;22:417-421.

52. Fitzgerald FT. The problem of obesity. Ann Rev Med 1981;32:221-231.

53. Flint MM, Drinkwater BL, Wells CL, Horvath SM. Validity of estimating body fat of females: effect of age and fitness. Hum Biol 1977;49(4):559-572.

54. Fomon SJ, Haschke F, Ziegler E, Nelson S. Body composition of reference children from birth to age 10 years. Am J Clin Nutr 1982;35:1169-1175.

55. Forbes GB. Relation of lean body mass to height in children and adolescents. Pediat Res 1972;6:32-37.

56. Forbes GB. Lean body mass and fat in obese children. Pediatrics 1964;34:308-314.

57. Forbes GB, Brown MR, Griffiths HJ. Arm muscle plus bone area: anthropometry and CAT scan compared. Am J Clin Nutr 1988;47:929-931.

58. Forbes GB, Amirhakimi GH. Skinfold thickness and body fat in children. Hum Biol 1970;42:401-418.

59. Forbes GB, Lewis A. Total sodium potassium and chloride in adult man. Journ Clin Invest 1956;35:596-600.

60. Forsyth R, Plyley MJ, Shepherd RJ. Residual volume as a tool in body fat prediction. Ann Nutr Metab 1988;32:62-67.

61. Franklin BA, Rubenfire M. Losing weight through exercise. JAMA 1980;244(4):377-379.

62. Frisancho AR, Flegel PN. Relative merits of old and new indices of body mass with reference to skinfold thickness. Am J Clin Nutr 1982;36:697-699.

63. Garfinkel L. Overweight and Cancer. Ann Intern Med 1985;103(6 pt 2):1034-1036.

64. Garn SM, Leonard WR, Hawthorne VM. Reply to letter from Miccozi & Albanes. Am J Clin Nutr 1987;46:377.

65. Garn SM, Leonard WR, Hawthorne VM. Three limitations of the body mass index. Am J Clin Nutr 1986;44:996-997.

66. Garn SM, Clark DC. Nutrition, growth, development and maturation: findings from the Ten-State Nutrition Survey of 1968-1970. Pediatrics 1975;56(2):306-319.

67. Garrison RJ, Castelli WP. Weight and thirty-year mortality of men in the Framingham study. Ann Intern Med 1985;103(6 pt 2):1006-1009.

68. Garrow J. Obesity - Public health enemy No. 2? Health Educ J 1986;45(1):57-58.

69. Garrow J, Webster J. Quetelet's Index (W/H^2) as a measure of fatness. Int J Obesity 1985;9:147-153.

70. Gordon T, Doyle J. Weight and mortality in men: the Albany Study. Int J Epidemiology 1988;17(1):77-81.

71. Gwinup G, Chelvam R, Steinberg T. Thickness of subcutaneous fat and activity of underlying muscles. Ann Intern Med 1971;74:408-411.

72. Hager A. Estimation of body fat in infants, children and adolescents. In: Bonnet P, ed, Adipose Tissue in Childhood. Boca Raton, Florida: CRC Press, 1981:49-56.

73. Hager A. Nutritional problems in adolescence - obesity. Nutr Rev 1981;39(2):89-95.

74. Hamill PV, Drizd TA, Johnson CL, Reed RB, Roche AF, Moore WM. Physical growth: National Center for Health Statistics percentiles. Am J Clin Nutr 1979;32:607-629.

75. Hammar S, Campbell M, Campbell V, Moores N, Sareen C, Gareis F, Lucas B. An interdisciplinary study of adolescent obesity. J Pediatrics 1972;80(3):373-383.

76. Harris T, Cook F, Garrison R, Higgins M, Kannel W, Goldman L. Body mass index and mortality among nonsmoking older persons. The Framingham Heart Study. JAMA 1988;259(10):1520-1524.

77. Harrison GE. The measurement of total body electrical conductivity. Hum Biol 1987;59(2):311-317.

78. Harrison GG. Height-weight tables. Ann Intern Med 1985;103(6 pt 2):989-994.

79. Harsha DW, Smoak CG, Nicklas TA, Webber LS, Berenson GS. Tracking of Body Composition Variables. Pediatrics. 1987;80(5 pt 2):779-783.

80. Haschke F. Body composition of adolescent males. Acta Paediat Scand Suppl 1983;307:3-23.

81. Haschke F, Fomon SJ, Ziegler EE. Body composition of a nine-year-old reference boy. Pediatr Res 1981;15:847-849.

82. Hayes PA, Sowood PJ, Belyavin A, Cohen JB, Smith FW. Sub-cutaneous fat thickness measured by magnetic resonance imaging, ultrasound, and calipers. Med Sci Sport Exer 1988;20(3):303-309.

83. Health Promotion Directorate, Health Services and Promotion Branch, Health and Welfare Canada. Canadian Guidelines for Healthy Weights. Ottawa: Minister of National Health and Welfare, 1988.

84. Healthy Living Institute. How to measure obesity. 10. Infrared interactance. Int Obesity Newsletter 1988;2(12):6-7.

85. Hodgdon JA, Fitzgerald PI. Validity of impedance predictions at various levels of fatness. Hum Biol 1987;59(2):281-298.

86. Jackson AS. Research design and analysis of data procedures for predicting body denisty. Med Sci Sport Exer 1984;16(6):616-620.

87. Jackson AS, Pollock ML, Ward A. Generalized equations for predicting body density of women. Med Sci Sport Exer 1980;12:175-182.

88. Jackson AS, Pollock ML. Generalized equations for predicting body density of men. Br J Nutr 1978;40:497-504.

89. Jackson AS, Pollock ML. Factor analysis and multivariate scaling of anthropometric variables for the assessment of body composition. Med Sci Sport Exer 1976;8(3):196-203.

90. Jequier E. Energy, obesity, and body wieght standards. Am J Clin Nutr 1987;45:1035-1047.

91. Johnston FE. Validity of triceps skinfold and relative weight as measures of adolescent obesity. J Adolesc Health Care 1985;6(3):185-190.

92. Johnston FE. Health implications of childhood obesity. Ann Intern Med 1985;103(6 pt 2):1068-1072.

93. Johnston FE. Relationships between body composition and anthropometry. Hum Biol 1982;54(2):221-245.

94. Johnston FE, Wadden TA, Stunkard AJ, Pena M, Wang J, Pierson R, Van Itallie T. Body fat deposition in adult obese women. I. Patterns of fat distribution. Am J Clin Nutr 1988;47:225-228.

95. Jones PR, Hunt MJ, Brown TP, Norgan NG. Waist-hip circumference ratio and its relation to age and overweight in British men. Hum Nutr:Clin Nutr 1986;40C:239-247.

96. Katch F, Michael ED, Horvath SM. Estimation of body volume by underwater weighing: description of a simple method. J Appl Physiol 1967;23(5):811-813.

97. Katch VL. A reaction to laboratory methodology. Med Sci Sport Exer 1984;16(6):604-605.

98. Keys A. Overweight, obesity, coronary heart disease and mortality. Nutr Rev 1980;38(9):297-307.

99. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. J Chron Dis 1972;25:329-343.

100. Keys A, Brozek J. Body fat in adult man. Physical Rev — 1953;33:245-325.

101. Khaled MA, McCutcheon MJ, Reddy S, Pearman PL, Hunter GR, Weinsier RL. Electrical impedance in assessing human body composition: the BIA method. Am J Clin Nutr 1988;47:789-792.

102. Korsch B. Childhood obesity. J Pediatrics 1986;109(2):299-300.

103. Kramer MS, Barr RG, Leduc DG, Boisjoly C, Pless IB. Infant determinants of childhood weight and adiposity. J Pediatrics 1985;107(1):104-107.

104. Kuczmarski RJ, Fanelli MT, Koch GG. Ultrasonic assessment of body composition in obese adults: overcoming the limitations of the skinfold caliper. Am J Clin Nutr 1987;45:717-724.

105. Lew EA. Mortality and weight: insured lives and the American Cancer Society Studies. Ann Intern Med 1985;103(6 pt 2):1024-1029.

106. Lloyd RD, Mays CW. A model for human body composition by total body counting. Hum Biol 1987;59(1):7-30.

107. Lohman TG. The use of skinfold to estimate body fatness on children and youth. JOPERD 1987;Nov-Dec:98-102.

108. Lohman TG. Applicability of body composition techniques and constants for children and youths. In: Pandolf K, ed. Exercise and Sports Sciences Reviews 1986;14:325-357.

109. Lohman TG. Research progress in the validation of laboratory methods of assessing body composition. Med Sci Sport Exer 1984;16(6):596-603.

110. Lohman TG. Skinfolds and body density and their relation to body fatness: a review. Hum Biol 1981;53(2):181-225.

111. Lohman TG, Slaughter MH, Boileau RA, Bunt J, Lussier L. Bone mineral measurements and their relation to body density in children, youth and adults. Hum Biol 1984;56(4):667-679.

112. Lukaski HC. Methods for the assessment of human body composition: traditional and new. Am J Clin Nutr 1987;46:537-556.

113. Malina RM. Bioelectric methods for estimating body composition: an overview and discussion. Hum Biol 1987;59(2):329-335.

114. Malina RM. Comments on clinical methods of assessing body composition. Med Sci Sport Exer 1984;16(6):614-615.

115. Manson JE, Stampfer MJ, Hennekens CH, Willett WC. Body weight and longevity. A reassessment. JAMA 1987;257(3):353-358.

116. Marshall WA. The relationship of puberty to other maturity indicators and body composition in man. J Reprod Fert 1978;52:437-443.

117. Martin AD. An anatomical basis for assessing human body composition: evidence from 25 dissections. PhD thesis, Simon Fraser University, Burnaby, Canada; as cited in Roche AF. Some aspects of the criterion methods for the measurement of body composition. Hum Biol 1987;59(2):209-220.

118. Martin AD, Drinkwater DP, Clarys JP, Ross WD. The inconstancy of the fat free mass: a reappraisal with implications for densitometry. In: T Reilly, J Watkins, J Borns (Eds.). Kinanthropometry III. London:E. & F.N. Spon, 1986, p. 92-97.

119. McLaren DS. Three limitations of the body mass index. Am J Clin Nutr 1987;46:121.

120. Mellin L. Dietitian's role in the management of obesity in adolescents. Directions in Applied Nutrition 1987;1(9):1,7-8.

121. Mendez J, Lukaski HC, Buskirk ER. Fat-free mass as a function of max O_2 consumption and 24-hour urinary creatinine, and 3-methylhistidine excretion. Am J Clin Nutr 1984;39:710-715.

122. Meneely GR, Ball OT, Kory RC, Callaway JJ, Merrill JM, Mabe RE, Roehm DC, Kaltreider NL. A simplified closed circuit helium dilution method for the determination of the residual volume of the lungs. Am J Med 1960;28:824-831.

123. Micozzi MS, Albanes D. Three limitations of the body mass index (letter to the editor). Am J Clin Nutr 1987;46:376.

124. Micozzi MS, Albanes D, Jones DY, Chumlea WC. Correlations of body mass indices with weight, stature, and body composition in men and women in NHANES I and II. Am J Clin Nutr 1986;44:725-731.

125. Morbidity and Mortality Weekly Report (1987;37(1)). Prevalence of overweight in selected states - behavioural risk factor surveillance, 1986. JAMA 1988;259(6);797-798.

126. Moynihan CM, Stark O, Peckham CS. Obesity in 16-year-olds assessed by relative weight and doctors' rating. Int J Obesity 1986;10:3-10.

127. Mueller WH. The changes with age of the anatomical distribution of fat. Soc Sci & Med 1982;16:191-196.

128. Mueller WH, Wohleb JC. Anatomical distribution of subcutaneous fat and its description by multivariate methods: how valid are principal components? Am J Phys Anthropol 1981;54:25-35.

129. Mueller WH, Stallones L. Anatomical distribution of subcutaneous fat: skinfold site choice and construction of indices. Hum Biol 1981;53(3):321-335.

130. Mueller WH, Reid RM. A multivariate analysis of fatness and relative fat patterning. Am J Phys Anthrop 1979;50:199-208.

131. National Institutes of Health Consensus Development Panel on the Health Implications of Obesity. Health implications of obesity. National Institutes of Health Consensus Development Conference Statement. Ann Intern Med 1985;103(6 pt 2):1073-1077.

132. Nash HL. Body fat measurement : weighing the pros and cons of electrical impedance. Phys & SportsMed 1985;13(11):124-128.

133. Norman GR, Streiner DL. PDQ Statistics. Toronto: B.C. Decker Inc., 1986.

134. Oppliger RA, Looney MA, Tipton CM. Reliability of hydrostatic weighing and skinfold measurements of body composition using a generalizability study. Hum Biol 1987;59(1):77-96.

135. Paige DM. Obesity in childhood and adolescence. Special problems in diagnosis and treatment. Postgrad Med 1986;79(1):233-245.

136. Parizkova J. Growth and growth velocity of lean body mass and fat in adolescent boys. Pediat Res 1976;10:647-650.

137. Pereira L, White F. Prevalence and health consequences of obesity. NIN Rapport 1987;2(2):6-7.

138. Pollock ML, Jackson AS. Research progress in validation of clinical methods of assessing body composition. Med Sci Sport Exer 1984;16(6):606-613.

139. Poskitt EM. Management of obesity. Arch Disease in Childhood 1987;62:305-310.

140. Presta E, Casullo AM, Costa R, Slonim A, Van Itallie TB. Body composition in adolescents: estimation of total body electrical conductivity. J Appl Physiol 1987;63(3):937-941.

141. Richmond D, Blyler E, Linscheio T. The obese child. Am Fam Phys 1983;28:129-134.

142. Riumulla J, Durnin JV. Changes in body composition in adolescent boys. Eur J Clin Nutr 1988;42:107-112.

143. Robinson N, Lankford F, Gorman M. TWIGS - a weight-control program designed for children. J Am Dietit Assoc 1987;87(4):500-502.

144. Roche AF. Some aspects of the criterion methods for the measurement of body composition. Hum Biol 1987;59(2):209-220.

145. Roche AF. Research progress in the field of body composition. Med Sci Sport Exer 1984;16(6):579-583.

146. Roche AF, Seirvogel RM, Chumlea WM, Webb P. Grading body fatness from limited anthropometric data. Am J Clin Nutr 1981;34:2831-2838.

147. Rolland-Cachera MF, Sempe M, Guilloud-Bataille M, Patois E, Pequinot-Guggenbuhl F, Fautrad V. Adiposity indices in children. Am J Clin Nutr 1982;36:178-184.

148. Rookus MA, Burema J, Deurenberg P. Van Der Wiel-Wetzels W. The impact of adjustment of a weight-height index (W/H^2) for frame size on the prediction of body fatness. Br J Nutr 1985;54:335-342.

149. Schwartz MZ. Childhood obesity. Surg Clin of North Am 1979;59(6):995-1006.

150. Segal KR, Van Loan M, Fitzgerald PI, Hodgdon JA, Van Itallie TB. Lean body mass estimation by bioelectrical impedance analysis: a four-site cross-validation study. Am J Clin Nutr 1988;47:7-14.

151. Seidell JC, Oosterlee A, Deurenberg P, Hautyast J, Ruijs J. Abdominal fat depots measured with computed tomography: effects of degree of obesity, sex, and age. Eur J Clin Nutr 1988;42:805-815.

152. Seltzer C, Mayer J. An effective weight control program in a public school system. Am J Publ Health 1970;60(4):679-689.

153. Shear CL, Freedman DS, Burke GL, Harsha DW, Berenson GS. Body fat patterning and blood pressure in children and young adults. The Bogalusa Heart Study. Hypertension 1987;9:236-244.

154. Simopoulos AP. Characteristics of obesity: an overview. Ann NY Aca Sci 1987;499:4-13.

155. Simopoulos AP. The health implications of overweight and obesity. Nutr Reviews 1985;43(2):33-40.

156. Simopoulos AP, Van Itallie TB. Body weight, health, and longevity. Ann Intern Med 1984;100:285-295.

157. Sinning WE. Body composition assessment of college wrestlers. Med Sci Sport Exer 1974;4:139-145.

158. Siri WE. The gross composition of the body. Adv Biol Med Phys 1956;4:239-280.

159. Ross WD, Crawford SM, Kerr DA, Ward R, Bailey DA, Mirwald RM. Relationship of the body mass index with skinfolds, girths, and bone breadths in Canadian men and women aged 20-70 years. Am J Phys Anthropol 1988;77:169-173.

160. Ross WD, Martin AD, Ward R. Body composition and aging: theoretical and methodological implications. Coll Antropol 1987;11(1):15-44.

161. Slaughter MH, Lohman TG, Boileau RA, Stillman RJ, Van Loan M, Horswill CA, Wilmore JH. Influence of maturation on relationship of skinfolds to body density: a cross-sectional study. Hum Biol 1984;56(4):681-689.

162. Smith JF, Bishop PA. Hydrostatic weighing at residual volume and total lung capacity. J Sports Med 1987;27:17-20.

163. Smith T, Hesp R, Mackenzie J. Total body potassium calibrations for normal and obese subjects to two types of whole body counter. Phys Med Biol 1979;24:171-175.

164. Stallones R. Epidemiologic studies of obesity. Ann Intern Med 1985;103(6 pt 2):1003-1005.

165. Stark O, Lloyd JK. Some aspects of obesity in childhood. Postgrad Med Journal 1986;62:87-92.

166. Steen B. Body composition and aging. Nutr Rev 1988;46(2):45-51.

167. Sterner TG, EJ Burke. Body fat assessment: a comparison of visual estimation and skinfold techniques. Phys & Sportsmed 1986;14(4):101-107.

168. Szeluga DJ, Stuart RK, Utermohlen V, Santos GW. Nutritional assessment by isotope dilution analysis of body composition. Am J Clin Nutr 1984;40:847-854.

169. Tabachnick BG, Fidell LS. Using Multivariate Statistics. New York: Harper & Row, Publishers, 1983.

170. Tanner JM. Issues and advances in adolescent growth and development. J Adolesc Health Care 1987;8:470-478.

171. Thorland WG, Johnson GO, Tharp GD, Housh TJ, Cisar CJ. Estimation of body density in adolescent athletes. Hum Biol 1984;56(3):439-448.

172. Van Itallie TB. Health implications of overweight and obesity in the United States. Ann Intern Med 1985;103(6 pt 2):983-988.

173. Van Loan M, Mayclin P. Bioelectrical impedance analysis: is it a reliable estimator of lean body mass and total body water? Hum Biol 1987;59(2):299-309.

174. Verbrugge LM, Wingard DL. Sex differentials in health and mortality. Health Matrix 1987;5(2):3-19.

175. Vignolo M, Naselli A, Di Battista E, Mostert M, Aicardi G. Growth and development in simple obesity. Eur J Pediatr 1988;147:242-244.

176. Wadden TA, Stunkard AJ, Johnston FE, Wang J, Pierson R, Van Itallie T, Costello E, Pena M. Body fat deposition in adult obese women. II Changes in fat distribution accompanying weight reduction. Am J Clin Nutr 1988;47:229-234.

177. Weltman A, Levine S, Seip R, Vutran Z. Accurate assessment of body composition in obese females. Am J Clin Nutr 1988;48:1179-1183.

178. Weltman A, Janney C, Huber R, Rians CB, Katch F. Comparison of hydrostatic weighing at residual volume and total lung capacity in pre-pubertal males. Hum Biol 1987;59(1):51-57.

179. Weltman A, Katch VL. A non population-specific method for predicting total body volume and percent fat. Hum Biol 1978;50(2):151-158.

180. Wilmore JH, Buskirk ER, Digirolamo M, Lohman TG. Body composition. A round table. Phys & Sportsmed 1986;14(3):144-162.

181. Wilmore JH. A reaction to the manuscripts of Roche and Buskirk. Med Sci Sport Exer 1984;16(6):594-595.

182. Wong WW, Cochran WJ, Klish WJ, Smith EO, Lee LS, Fiorotto ML, Klein PD. Body fat in normal adults estimated by oxygen-18 and deuterium dilution and by anthropometry: a comparison. Eur J Clin Nutr 1988;42:233-242.

183. Woodward HQ, White DR. The composition of body tissues. Br J Radiology 1986;59:1209-1219.

184. Young CM, Sipin SS, Roe D. Body composition of pre-adolescent and adolescent girls. J Amer Dietit Assoc 1968;53:25-31.

185. Zack PM, Harlan WR, Leaverton PE, Coroni-Huntley J. A longitudinal study of body fatness in childhood and adolescence. J Pediatrics 1979;95(1):126-130.